The Asset Demand of U.S. Households

Xavier Gabaix Ralph S.J. Koijen Federico Mainardi Sangmin S. Oh Motohiro Yogo*

August 26, 2022 – Preliminary and incomplete

Abstract

We use new monthly security-level data on portfolio holdings and flows of U.S. households to estimate asset demand across asset classes and individual securities. Our data features broad coverage across the wealth distribution – including ultra-high-net-worth (UHNW) households – and spans multiple asset classes, covering both public and private assets. We first develop a factor model to explain households' rebalancing across asset classes and find that three factors, with clear economic interpretations, explain 65% of the variation. Across households, we find that during equity downturns UHNW households buy risky assets and in particular equities, and thus mitigate market fluctuations while less wealthy households sell and thus amplify market fluctuations. We then estimate household-level demand curves for the cross-section of U.S. equities. We find that demand elasticities decline in wealth, and we confirm earlier findings that they are very low on average. Lastly, we develop a new framework to estimate demand curves across asset classes. While nesting traditional models as a special case, our framework allows for inertia and a muted response of asset demand to fluctuations in asset prices. A calibrated extension of the model can be used to understand the large response of risky asset prices to monetary policy shocks. In sum, we think that the framework and data we provide allow to propose a coherent picture of US households, including, quite uniquely, their rebalancing behavior across broad asset classes.

*xgabaix@fas.harvard.edu, ralph.koijen@chicagobooth.edu, fmainard@chicagobooth.edu, oh@chicagobooth.edu, myogo@princeton.edu. Gabaix thanks the Ferrante Fund for financial support. Koijen acknowledges financial support from the Center for Research in Security Prices at the University of Chicago Booth School of Business.

1 Introduction

We study the asset demand of households, including ultra-high-net-worth (UHNW) households, in the United States (US) across individual securities and asset classes. Households play a central role in modern asset pricing models, either by investing directly in financial markets or by allocating capital to intermediaries. Yet, the data available in the US are still quite limited (we discuss the related literature in detail below).

We use proprietary data from Addepar, a wealth management platform for investment advisors, to fill this gap. Addepar provides wealth managers with real-time portfolio information to guide investment decisions. The Addepar data contains security-level information on holdings, flows, and returns that aggregate to narrow asset classes (e.g., US equities or US investment-grade corporate bonds) and broad asset classes (e.g., equities, fixed income, and alternatives). We observe the data at a monthly frequency from January 2016 to August 2021.[1] The platform has been growing rapidly during our sample period, and the total assets (number of portfolios) in our data increase from $314 billion (22,683) to $2.23 trillion (227,261).

Our data have two important advantages. First, we have data on UHNW individuals, with well over a thousand households who own more than $100 million in assets. This group of households, which may be particularly relevant for asset prices, is typically underrepresented in other data sources. The broad coverage across the wealth distribution also allows us to extrapolate our estimates to construct demand curves for the "representative US household." Second, we have broad coverage across asset classes and at high frequencies. The asset classes covered in the data include public and private assets and are all disaggregated to security-level data. Such a broad perspective is not even available for most US institutions.

We provide four main contributions. First, we provide two methodological contributions to analyze flows and asset demand across asset classes. We first develop a descriptive factor model for flows to measure (i) how investors allocate capital to cash and to all risky asset classes combined and (ii) how investors reallocate capital across risky asset classes. Next, we develop a new methodology to identify and estimate demand curves across asset classes. This approach relies on economically-motivated restrictions on the covariance matrix of flows and returns, and can be applied in many settings where broad coverage of holdings is unavailable. Second, our empirical results show that the flow to risky assets (and particularly to equities) is pro-cyclical for less wealthy households (assets below $3 million) and counter-cyclical for wealthy households (assets above $10 million and in particular above $100 million). Wealthy households therefore stabilize fluctuations in risky asset markets, while less wealthy households contribute to markets' excess volatility. Third, the factor model identifies three key rebalancing factors that are economically meaningful. These factors capture bets on the long-term equity premium, the credit premium, the municipal bond premium,

[1] We receive updates of the data with a 6-month delay.

and the premium on US equities versus global equities. Fourth, we estimate a demand model for the cross-section of US equities and find that demand elasticities for individual stocks decline in wealth and – in line with earlier literature – are generally very low.

The paper is organized in three parts. In the first part of the paper, we study allocations and flows across asset classes. For allocations, we document three main facts. First, wealthier households allocate a larger fraction to alternatives, such as private equity and hedge funds, while reducing their allocations to fixed income and public equities. Second, among liquid risky asset classes, wealthier households allocate a larger fraction to municipal and tax exempt bonds by reducing their allocations to investment-grade corporate bonds and global equities. This wealth gradient can be explained by the tax benefits of municipal bonds.[2] Third, even though allocations vary across the wealth distribution, heterogeneity in wealth can only explain a small fraction of the variation in portfolio allocations across households. This highlights the importance of other factors beyond wealth in explaining households' portfolios, such as heterogeneous beliefs, risk preferences, and portfolio inertia.

Next, we study the flows to cash and liquid risky assets. The flow to liquid risky assets aggregates the flows across 14 asset classes of which US equities is the largest. This analysis reveals three new insights. First, the average flows to liquid risky assets and cash are strongly negatively correlated. In addition, while the average flow to risky assets is strongly positively correlated with equity returns, the dispersion in flows across investors is negatively correlated with returns. This implies that, on average, investors sell during economic downturns and that, at the same time, disagreement increases during those turbulent times. Second, we find that flows to cash are much more volatile compared to the average fraction invested in cash. This illustrates the dual role that cash plays in households' portfolios: on the one hand, cash is used as a safe asset, but, on the other hand, it is also used to buffer liquidity shocks that do not affect the allocations to liquid risky assets. Third, we estimate how the flow to liquid risky assets responds to aggregate stock returns across the wealth distribution. Quite strikingly, we find that the sensitivity declines sharply in wealth. In fact, the flows of households with assets over $100 million are insensitive to stock returns. This low sensitivity could be consistent with wealthy households being inert or it can reflect rebalancing within liquid risky assets from Treasuries to US equities, for instance. To separate these hypotheses, we estimate the sensitivity of flows to US equities to aggregate stock returns. We find that while less wealthy households act pro-cyclically, UHNW households buy equities during downturns and thus stabilize markets by providing elasticity. That said, the magnitudes are small, suggesting that the total elasticity that UHNW households provide is limited.

To complete the first part of the paper, we study the rebalancing across liquid risky asset classes once we account for the aggregate flows to liquid risky assets and cash. We introduce a simple

[2]We refer to Balloch and Richers (2021) for additional facts about allocations across the wealth distribution using Addepar data.

decomposition to estimate the main rebalancing directions in the data using principal components analysis (PCA). We find that the first three principal components explain approximately 65% of all rebalancing variation across the 14 asset classes. In addition, we show that the factor loadings form a long-short portfolio across asset classes and summarize the key rebalancing directions of households (although they may disagree on the direction and magnitude of the trade). The three factors carry intuitive economic interpretations. The first factor rebalances from US equities to long-duration fixed income such as US treasuries and agencies, municipal and tax exempt bonds, and US investment-grade corporate bonds. This factor therefore bets on the long-term equity risk premium. The second factor rebalances from US investment-grade bonds to US Treasuries, and thus represents a bet on the credit premium. The third factor combines two trades. The first trade rebalances from US Treasuries and US investment-grade corporate bonds to municipal and tax exempt bonds, while the second rebalances from global equities to US equities. The third factor thus bets on the premium for municipal bonds as well as global equities versus US equities. Taken together, this part of the paper summarizes key dimensions of rebalancing in terms of cash, the aggregate flow to liquid risky assets, and rebalancing flows across liquid risky assets that are informative for the design of macro-finance models with rich heterogeneity. In ongoing work, we are evaluating the effects of monetary policy shocks[3] on flows and portfolio rebalancing.

The first part of this paper is mostly descriptive. In the second and third parts, we develop a more structural model of the demand for assets, including across asset classes. In the second part of the paper, we estimate household-level demand curves for the cross-section of US equities. This analysis extends the recent literature on demand system asset pricing (Koijen and Yogo, 2019). The goal of this literature is to develop equilibrium asset pricing models that are consistent with observed asset prices, portfolio holdings, and firm characteristics. Typically, data on institutional holdings are used to estimate the models and the (aggregated) holdings of the household sector are computed as the difference between a firm's shares outstanding and the aggregate number of shares held by institutions. While this is the best practice when only institutional data are available, we make progress by using our data to estimate household-level asset demand curves.

This analysis reveals three main insights. First, as an introductory analysis, we document the link between an household's active share and wealth. Cremers and Petajisto (2009) show, and we confirm in our institutional data, that larger institutions are more passive and have a lower active share. For households, however, we find the opposite: wealthier households have a higher active share and are more active. Second, we estimate the demand model of Koijen and Yogo (2019) for individual households. A key finding of the earlier literature is that demand elasticities (for individual stocks) are much lower than those implied by theoretical models. We confirm this finding for households. Moreover, we document an important new fact that demand elasticities

[3]We follow Gürkaynak et al. (2005), Nakamura and Steinsson (2018), and Bauer and Swanson (2022) in constructing monetary policy shocks.

decline in wealth. For households with assets below $3 million, we estimate an average elasticity of 0.82, while it drops to 0.58 for households with more than $100 million in assets. Third, we find that demand elasticities drop at the onset of the COVID-19 pandemic during the first quarter of 2020, and recover thereafter. This pro-cyclicality of elasticities is consistent with liquidity evaporating in financial markets during times of stress. In ongoing work, we also explore how these estimates change if we account for the indirect holdings of households via mutual funds and exchange-traded funds (ETFs). This extension again leverages the strength of the unique data that we use.

In the third and final part of the paper, we estimate asset-class level demand curves for households. Estimating demand elasticities across asset classes is particularly challenging. Recently proposed methods rely on idiosyncratic demand shocks to one group of investors to generate exogenous variation in prices that can be used to estimate elasticities of other investors (Gabaix and Koijen, 2021; Koijen and Yogo, 2020). As we only observe portfolio holdings for a subset of households, and we do not observe the holdings of institutions at the same high frequency and across many asset classes, these recent methods cannot be directly used in this setting without additional data.

To overcome this problem, we propose a new method to identify demand curves by imposing additional structure that is grounded in economic theory. In particular, we show that in a large class of models, the own-price elasticity depends on one minus the portfolio weight and the cross-price elasticity depends on the negative of the portfolio share. We show that this structure imposes restrictions on the covariance matrix of flows and returns that can then be used to estimate flexible models of asset-class demand. In particular, our framework allows for inertia and for a muted response to prices and expected returns, consistent with the recent evidence on inelastic financial markets. At the same time, the framework nests the traditional mean-variance portfolio choice model as a special case. This new approach to asset demand estimation can be used in settings where incomplete data on holdings are available or where large granular shocks do not occur.

Finally, we show how the model can be extended to account for reaching for yield effects, and provide a simple calibration to show how the model can potentially reconcile the strong response of risky assets to monetary policy shocks (Bernanke and Kuttner, 2005).

In ongoing work, we are estimating the model to quantify the degree of inertia, the (in)sensitivity of asset-class level demand to risk premia, and the strength of the reaching for yield effect.

The paper proceeds as follows. In Section 2, we introduce the data, sample selection, and we provide summary statistics. We then study the allocation and flows across narrow asset classes in Section 3. In Section 4, we estimate household-level demand curves for the cross-section of US equities. We extend this analysis in Section 5 to estimate demand curves across narrow asset classes. We conclude in Section 6.

Related literature

Our paper mainly contributes to two literatures. First, our paper contributes to the literature on asset demand of households, including high-net-worth households. We summarize this literature in Table A1 in Appendix A. In this literature, one line of work has focused on the investment behavior of households in the US using proprietary data, mainly from large financial institutions or brokerage firms. Early examples include Barber and Odean (2000), who study the trading behavior of retail investors from 1991 to 1996 using data from a large discount broker, and Ameriks and Zeldes (2004) who analyze the equity share over the life cycle using data from the Survey of Consumer Finances and TIAA-CREF.

More recently, Giglio et al. (2021a) use data from a survey administered to Vanguard clients matched to administrative data on portfolio allocation to estimate the pass-through from beliefs to actions. Related to Giglio et al. (2021a) is Bender et al. (2022), who also use data from a survey administered through UBS to a sample of 2,484 affluent US investors. Egan et al. (2021) use data from BrightScope Beacon on portfolio allocations for a large sample of 401(k) plans. Parker et al. (2022) study asset allocation decisions over the life cycle using a proprietary dataset from a large US financial institution, highlighting the significant impact that target date funds have had in recent years. Other papers analyze the behavior of wealthy households by focusing on tax-return rather than proprietary data. Among others, Hoopes et al. (2016) uses administrative data from IRS at daily frequency between 2008 and 2009 to analyze the behavior of individual investors during the market turmoil at the beginning of the Great Financial Crisis.

Balloch and Richers (2021) also use data from Addepar to study how asset class allocations and investment returns vary across the wealth distribution. Our paper complements Balloch and Richers (2021) by studying how investors rebalance their portfolios across asset classes and by estimating demand curves, both at the level of asset classes and at the level of US equities.

Outside the US, there are detailed data available on household portfolio holdings in several Scandinavian countries and in India. For Sweden, Calvet et al. (2007b) and Calvet et al. (2009a) (among others) use government-collected disaggregated data about income and wealth for the universe of Swedish households for income tax and wealth tax purposes, focusing on the 1999 to 2002 period. Calvet et al. (2021) use Swedish administrative data over a longer sample period and estimate the cross-sectional distribution of structural pereference parameters in a rich life-cycle model of saving and portfolio choice. Similar data are available in Norway. For instance, Fagereng et al. (2020) examine heterogeneity in returns across the wealth distribution both within and across asset classes. Betermier et al. (2022) construct factors by sorting stocks based on characteristics of investors that own them, and find that they explain both variation in portfolio holdings and cross-sectional variation in stock returns. For India, Campbell et al. (2019) use data from National Securities Depository Limited which include equity positions directly held by 11.6 million households between 2002 and

2011.

Second, our paper contributes to the recent literature on demand system asset pricing (Koijen and Yogo, 2019; Gabaix and Koijen, 2021; Haddad et al., 2022; Bretscher et al., 2022). The goal in this literature is to jointly understand data on prices, portfolio holdings and flows, and firm characteristics or macro variables. A key finding that has emerged from this literature is that data is much more inelastic compared to standard theories, which can be important in explaining excess volatility in financial markets. Our main contribution is that we can study disaggregated data on US households, not only for US equities but also across a large number of asset classes.

1.1 Definitions and notation

We denote time by t and investors by i, $i = 1, ..., I$. We denote security-level asset holdings by a (e.g., Apple versus Google stock), which can be aggregated to narrow asset classes that we index by n (e.g., US equities or US Treasuries) or broad asset classes that we index b (e.g., equities or fixed income). We provide the precise definitions of asset classes in Section 2.2. To introduce the variable definitions, we use narrow asset classes to index variables, but the same definitions apply to individual securities and broad asset classes.

We denote assets by A_{int}, dollar flows by F_{int}, and dollar returns by $R^{\$}_{int}$. We also observe time-weighted returns in our data, which we denote by r_{int}.[4] The inter-period budget constraint is then given by

$$A_{int} = A_{in,t-1} + R^{\$}_{int} + F_{int}. \tag{1}$$

We denote the aggregate flow by $F_{it} \coloneqq \sum_n F_{int}$, aggregate assets by $A_{it} \coloneqq \sum_n A_{int}$, and the aggregate dollar return by $R^{\$}_{it} \coloneqq \sum_n R^{\$}_{int}$. We define portfolio weights as $\theta_{int} = \frac{A_{int}}{A_{it}}$.

We denote flows, expressed as a fraction of total assets, by $f_{int} = \frac{F_{int}}{A^{DH}_{i,t-1}}$, where $A^{DH}_{i,t-1} \coloneqq \frac{1}{2}(A_{it} - R^{\$}_{it} + A_{i,t-1}) = A_{i,t-1} + \frac{1}{2}F_{it}$. In this definition, $A_{it} - R^{\$}_{it}$ corresponds to end-of-period wealth in the absence of price effects. Our definition of flows is similar in spirit to Davis and Haltiwanger (1992), and leads to a more robust definition of flows when $A_{i,t-1} = 0$ or small.

For the total flow, we define

$$f_{it} = \frac{F_{it}}{A^{DH}_{i,t-1}} = \sum_n \frac{F_{int}}{A^{DH}_{i,t-1}} = \sum_n f_{int}, \tag{2}$$

which implies $f_{it} \in [-2, 2]$.

[4] We winsorize the monthly time-weighted return at -300% and 300% for each security before aggregating the returns at the level of narrow asset classes using value weights.

2 Data and summary statistics

2.1 Data sources

Addepar Our primary data source is Addepar. Addepar is a wealth management platform that specializes in data aggregation, analytics, and reporting for complex investment portfolios that include public and private assets. It provides asset owners and advisors an overview of their financial positions. When possible, Addepar directly receives data on holdings and flows from custodians. It then combines this with information on returns from various data sources.

Addepar works with over 400 financial advisors, family offices, and large financial institutions that manage data for over $3 trillion of assets on the company's platform, ranging from the affluent to the ultra-high-net-worth investor segments.

Our sample contains monthly data from January 2016 to August 2021. We receive monthly updates with a delay of six months to preserve the confidentiality of the data. Given our main focus on flows, we aggregate the data to quarterly observations, as it may take some time for households to rebalance their portfolios in response to new information.

Portfolios are the unit of observation in Addepar. The same household or family can have multiple portfolios. We cannot identify which portfolios are connected[5] and we will refer to portfolios as households.

Addepar imposes two additional screens to preserve the confidentiality of the data. First, households receive investment advice from financial advisors. Firms that make up more than 10% of all portfolios in a given month are removed. If a portfolio is once removed via this process, it will not appear in subsequent months. Second, Addepar removes concentrated positions that exceed $1 billion in equities or companies that can be traced back to show the identity of an investor. We do observe which accounts are affected by this screen.

FactSet, CRSP, and Compustat As a point of reference, we compare the portfolios of households to those of institutional investors based on data from Security and Exchange Commission (SEC) Form 13F. All institutional investment managers that exercise investment discretion on accounts holding Section 13(f) securities, exceeding $100 million in total market value, must file the form. The data are sourced from FactSet. We merge the holdings data with data on stock prices from CRSP and accounting data from Compustat. The data construction follows Koijen et al. (2021).

Tick Data We use high-frequency data on Eurodollar futures sourced from Tick Data to construct monetary policy shocks, following Gurkaynak, Sack, and Swanson (2005) and Nakamura

[5]For some observations, the holdings in multiple asset classes are identical, suggesting that those are multiple portfolios belonging to the same household of family.

and Steinsson (2018). We follow the data construction steps in Bauer and Swanson (2022) and we provide further details in Appendix B.

2.2 Asset class definitions and summary statistics

Our sample of Addepar data includes information on 227,261 distinct client portfolios from 2016.Q1 to 2021.Q3. Table 1 provides the definition of asset classes in the Addepar data. In Online Appendix A, we describe each of the asset classes in more detail.

Table 1: Asset class definitions
This table reports the asset class taxonomy. We categorize the narrow asset classes, which are indexed by n, into seven broad asset classes. The narrow asset classes are obtained from Addepar's internal classification.

Broad asset classes	Narrow asset classes
Cash & Cash Equivalents	Cash, Cash Equivalents
Fixed Income	US Municipals/Tax Exempt, US Treasuries and Agencies, US TIPS, US Investment Grade, US High Yield, US Bank Loans, International Developed Markets, Emerging Markets, Opportunistic, Other Fixed Income, Unknown Fixed Income
Equities	US Equities, Concentrated Equity Positions, Global Equities, Developed Markets - Americas, Developed Markets - EMEA, Developed Markets - Asia Pacific, Emerging & Frontier Markets, Other Equities, Unknown Equities
Mixed Allocation	Asset Allocation Vehicle, Held Away Accounts
Alternatives	Hedge Funds, Private Equity & Venture, Real Estate Funds, Concentrated Alts. Positions, Unknown Alts., Other Alts, Direct Private Companies, Direct Real Estate, Direct Loans
Non-Financial Assets	Collectibles and Other
Liability	Liability

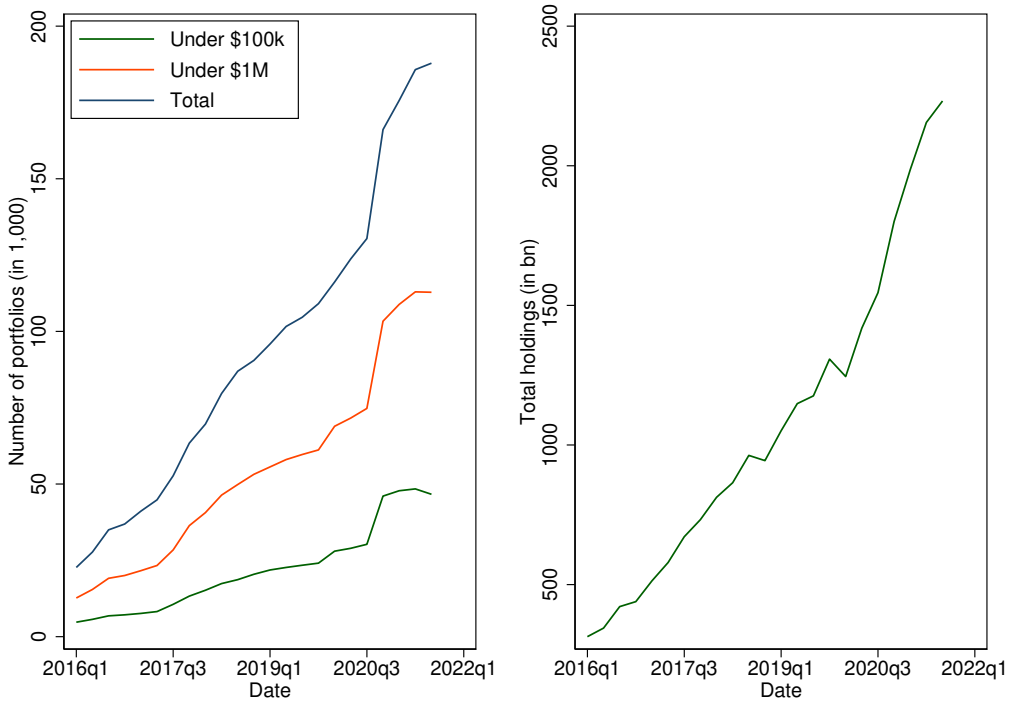
We make several adjustments to the classification of asset classes in the data. First, when holdings are classified as Unknown Equities and the investment sub type[6] is either a mutual fund or ETF, we relabel the asset class to US Equities when the Morningstar classification is U.S. Equity or to Global Equities when the Morningstar Classification is International Equity.

Second, we merge Unknown Equities into Other Equities and we do the same for fixed income. Third, we combine all Developed Markets (Americas, Asia Pacific, and EMEA) into a single equity category Development Markets. Lastly, we combine Cash and Cash Equivalents into a single narrow asset class Cash.

We define liquid and illiquid asset classes in our analysis below. Using the definitions in Table 1, the liquid asset classes include Developed Markets, Emerging & Frontier Markets, Emerging

[6]Investment type and subtype provide additional detail on the nature of the asset class.

Figure 1: Number of portfolios and total assets
In the left panel, we plot the total number of portfolios, the number of portfolios that are smaller than $1 million, and the number of portfolios that are smaller than $100k. In the right panel, we plot the total value of assets in our sample. The sample period is from January 2016 to August 2021.



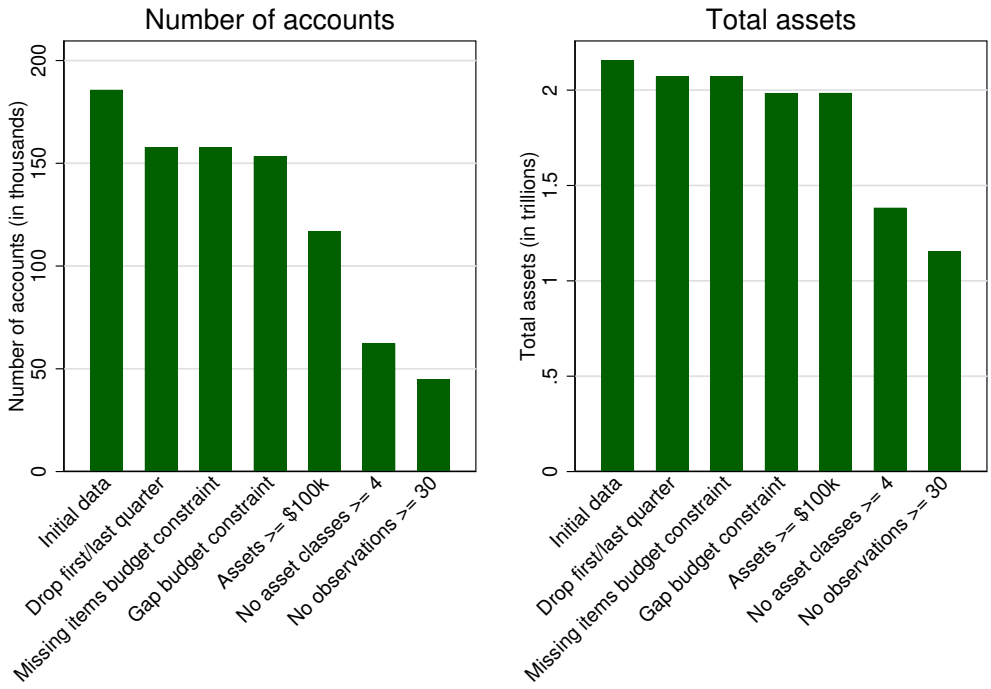
Markets, Global Equities, International Developed Markets, Opportunistic, US Equities, US High Yield, US Investment Grade Credit, US Municipals/Tax Exempt, US Tips, US Treasuries and Agencies, Other Equities, Other Fixed Income. We analyze cash separately for reasons we discuss below. All other narrow asset classes in Table 1 are classified as illiquid.

In Figure 1, we summarize the number of portfolios in the left panel and households' total assets on the platform in the right panel before imposing any screens. The number of portfolios grows from 22,683 in 2016.Q1 to 187,869 in 2021.Q3, and there are 227,261 unique portfolios in the sample. The sharp increase in the number of portfolios reflects the growth of the Addepar platform during our sample period. Households' total assets grow from $314 billion to $2.23 trillion during the same period.

2.3 Sample selection

We impose a series of sample selection screens in constructing our final sample. The sample selection screens ensure that the households are active in multiple asset classes so that we have broad coverage

Figure 2: The impact of sample selection screens on the number of portfolios and total assets
This figure summarizes the impact of the sample selection screens discussed in Section 2.3. In the left panel, we show the impact on the number of accounts. In the right panel, we show the impact on the total assets covered in our sample. The results are presented for 2021.Q2, the last complete quarter of our sample.



for the empirical analysis. Also, by imposing restrictions on the number of asset classes, we reduce the possibility that only part of a households' assets are covered on the Addepar platform. The screens also remove infrequent data errors. We will first discuss each of the screens. We then conclude this section by summarizing the impact of these screens on the coverage of our final sample.

We start by removing the quarter in which a household is onboarded onto the platform as flows tend to be more volatile during this period (for instance, as the beginning-of-period assets are unknown for some or all of the asset classses). We remove the last quarter that we observe a given household for the same reason.[7]

Second, we remove household-quarter observations when an item from the budget constraint is missing – that is, the starting value, $A_{in,t-1}$, the ending value, A_{int}, the flow, F_{int}, or the dollar return, $R^{\$}_{int}$. Third, we remove household-quarter observations where the budget constraint does not hold for at least one of the liquid narrow asset classes.[8] Fourth, for a small fraction of observations,

[7]It is uncommon for households to leave the platform during our sample period.

[8]We allow for a small margin of error of $1,000 or 0.5% of the average (absolute value) of the ending and starting

the starting value and ending value coincide. While this can happen for cash accounts, this is unlikely to be correct for risky assets. Therefore, we set returns and flows to zero for the liquid narrow asset classes that are not "Cash." This leads to an adjustment in 1.59% of all narrow asset class-quarter observations.[9]

Fifth, we drop household-quarter observations with fewer than $100k in assets (across both liquid and illiquid asset classes as well as cash). This screen also mitigates the concern that we capture only part of a household's assets. Sixth, we restrict to households with positive assets in the beginning or at the end of the period in at least four liquid asset classes. As we are interested in measuring rebalancing across asset classes, we focus on households who are active across multiple liquid asset classes. Lastly, we include a household only when there are more than 30 observations across all liquid asset classes and quarters. This screen ensures that we can estimate household-level factor loadings.

We summarize the impact of each of the screens in Figure 2 for the second quarter of 2021, which is the last complete quarter of our sample. In the left panel we report the total number of assets and in the right panel we report the total assets covered.

The sample selection screens that have a noticeable impact on the size of the sample are: first, to remove the onboarding quarter; second, to impose a size constraint; third, to require positive positions in at least four asset classes; and finally to require at least 30 observations across quarters and liquid asset classes. As wealthier households are more likely to satisfy these screens, the impact is larger in terms of the number of portfolios compared to total assets.

Based on a household's total wealth, we assign it to one of five wealth groups in a given quarter: A_{it} <\$3m, A_{it} ∈[\$3m, \$10m), A_{it} ∈[\$10m, \$30m), A_{it} ∈[\$30m, \$100m), and A_{it} ≥\$100m. We conclude our sample construction by winsorizing the flows at the 2.5% and 97.5% percentiles by narrow asset class and quarter, and balancing the panel in terms of holdings (across liquid and illiquid asset classes as well as cash) and flows (across liquid asset classes as well as cash).

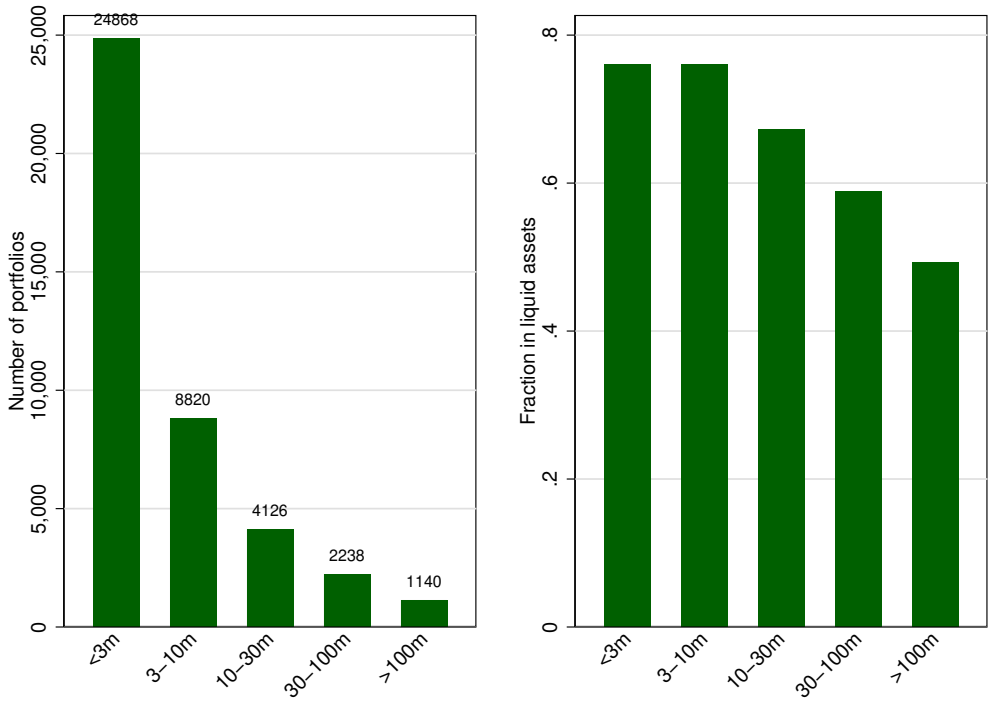
3 Asset demand across asset classes

We lead off our analysis by studying households' asset demand across asset classes. We proceed in three steps. First, we report a series of summary statistics regarding households' portfolio holdings in Section 3.1. These results complement the findings in Balloch and Richers (2021). We also report basic statistics on flows across broad asset classes. Second, we study the flow to risky liquid assets

value.

[9] In those cases, we often observe that the flow is the negative of the dollar returns. The reason is that the system has additional information about either the return or the flow, and completes the missing items in those instances to ensure that the budget constraint holds. Alternatively, we can drop those observations. However, as we balance the panel below, this alternative data construction step would be equivalent to setting those flows to zero and mis-measuring the level of assets.

Figure 3: Number of portfolios and the fraction invested in liquid assets by wealth group
In the left panel, we plot the number of portfolios in each of the five wealth groups. In the right panel, we plot the average fraction invested in liquid risky assets. The results are presented for 2019.Q4.



and cash in Section 3.2. Third, we provide a new framework to measure rebalancing across risky liquid asset classes and implement it in Section 3.3.

3.1 Summary statistics on portfolio holdings and flows

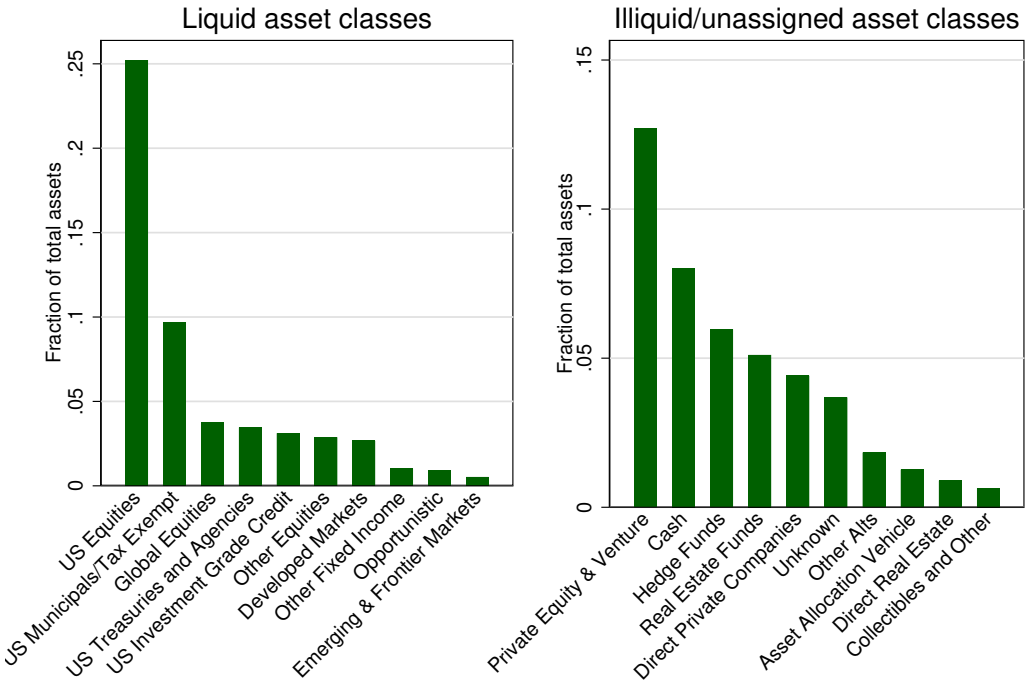
We first provide basic summary statistics on portfolio holdings across broad and narrow asset classes. We select a quarter in the middle of the sample, 2019.Q4, to present the results.

We plot the total number of portfolios in each of the wealth groups in the left panel of Figure 3. While the number of portfolios naturally declines in wealth, there are still 1,140 portfolios in our sample of more than $100 million in assets. We plot the fraction of total assets invested in liquid risky asset classes in 2019.Q4 in the right panel. As expected, wealthier households allocate a larger fraction to illiquid asset classes such as hedge funds, private equity, and other alternatives. We explore this pattern in more detail below.

In Figure 4, we plot the average portfolio shares across investors in 2019.Q4 for the 10 largest liquid risky asset classes (left panel) and the 10 largest illiquid asset classes (right panel).[10] Among

[10]We treat cash separately (that is, we do not classify it as part of the liquid risky assets and illiquid assets) for

Figure 4: Fraction invested in narrow asset classes
In the left panel, we plot the average portfolio shares in the largest 10 liquid risky asset classes. In the right panel, we plot the portfolio shares for the illiquid asset classes as well as cash. The results are presented for 2019.Q4.



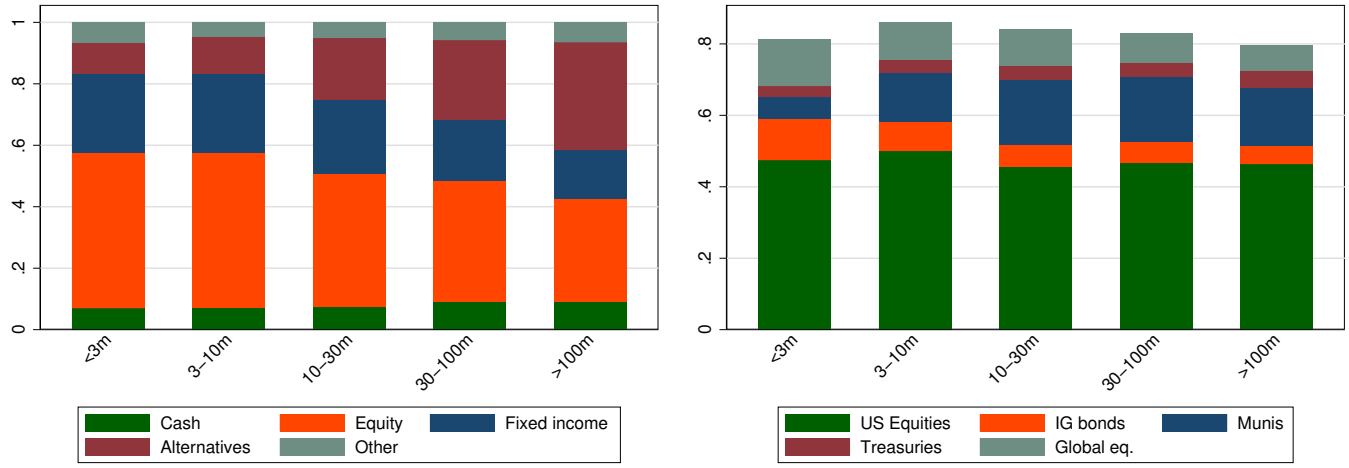
liquid asset classes, US equities is the dominant asset class, followed by municipal bonds, global equities, US Treasuries, and US investment grade bonds. Among illiquid asset classes, the dominant asset class is private equity and venture capital, followed by hedge funds, and real estate funds.

We summarize the fraction invested in broad asset classes by wealth group in 2019.Q4 in the left panel of Figure 5. In line with the right panel of Figure 3, wealthier households allocate a larger fraction to alternatives, while reducing their portfolio shares in public equity and fixed income. The fractions invested in cash and other assets is stable across the wealth distribution.

We plot the portfolio shares invested in the five largest liquid risky asset classes across the wealth distribution in the right panel of Figure 5, which are US equities, municipal and tax-exempt bonds, Treasuries, US investment-grade bonds, and global equities. These five asset classes account for approximately 80% of all assets invested in liquid risky assets. While the shares are fairly stable, the fraction invested in municipal bonds increases with wealth, at the expense of US investment grade bonds and global equities. This pattern may be driven by the tax benefits provided by municipal bonds.

reasons that we discuss in Section 3.2. In Figure 4, we report the average share in cash in the right panel, having noted that we do not treat it as an illiquid asset class.

Figure 5: Fractions invested in broad and narrow asset classes by wealth group
In the left panel, we plot the average fractions invested in broad asset classes (Cash, Equity, Fixed income, Alternatives, Other). In the right panel, we plot the average fractions invested in the five largest liquid risky asset classes (US Equities, US Investment-grade bonds, Municipal and tax-exempt bonds, Treasuries, and Global equities). The results are presented for 2019.Q4.



The previous figures illustrate important wealth gradients in asset allocations to different asset classes. That said, wealth cannot explain all or even most of the heterogeneity in observed portfolio holdings. To illustrate this, we estimate the following regression:

$$\theta_{int} = a_{0n} + a_{1n} \ln A_{it} + e_{int}, \tag{3}$$

at the level of broad and narrow asset classes and we record the R^2 value in Figure 6. We also report the standard deviation of θ_{int} to summarize the heterogeneity in portfolio holdings in a simple way.

We focus on broad asset classes in the left panel and on the five largest liquid risky asset classes in the right panel. Across the board, we find that the R^2 values are low, as is commonly observed in the household finance literature. The fraction invested in municipal bonds is best explained by wealth with an R^2 value close to 10%. This implies that there are other determinants of households' portfolios such as differences in beliefs, perceptions of risk, risk preferences, et cetera that are more important in explaining heterogeneity in portfolio shares.

To conclude this section, we summarize the pattern in flows to broad asset classes. We will explore the flows to liquid risky asset classes in great detail in the remainder of the paper. As reallocating capital across illiquid asset classes such as private equity and hedge funds is typically not feasible at high frequencies, we focus on the long-term trends during our sample.

In Figure 7, we plot the cumulative flows to broad asset classes during our sample. During this 5-year period, the cumulative flows have been positive for fixed income and equities, and negative for cash (which includes money market funds). One potential interpretation is that households

Figure 6: Heterogeneity in portfolio shares that cannot be explained by wealth
The orange bars correspond to the standard deviation of portfolio shares in broad asset classes (left panel) and the largest five narrow asset classes (right panel). The green bars correspond to the R-squared of a regression of portfolio weights on log wealth (see (3)). The results are presented for 2019.Q4.

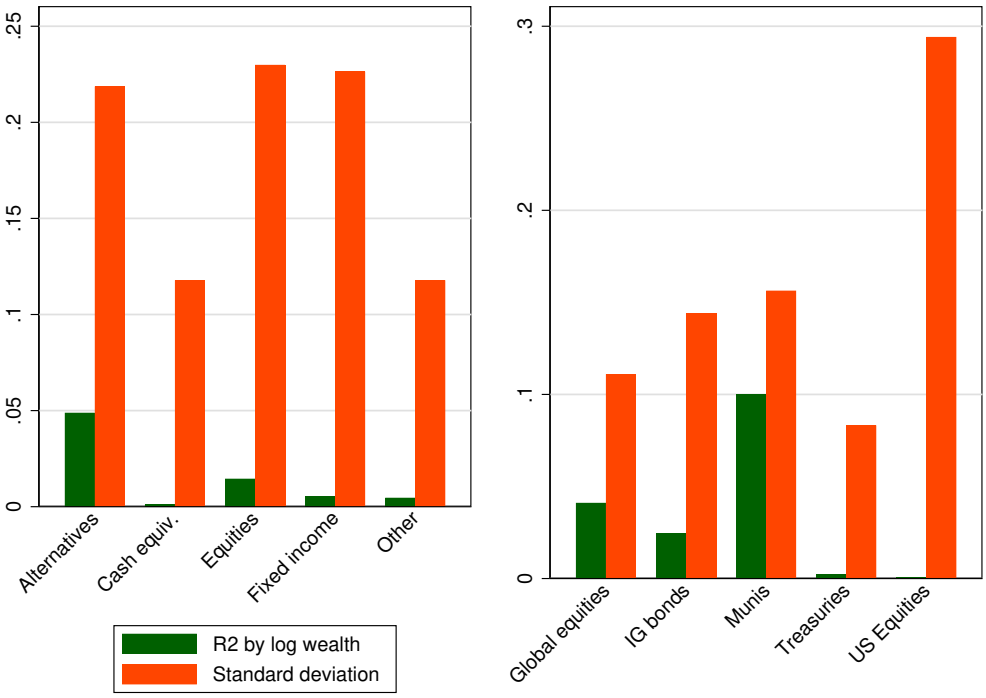
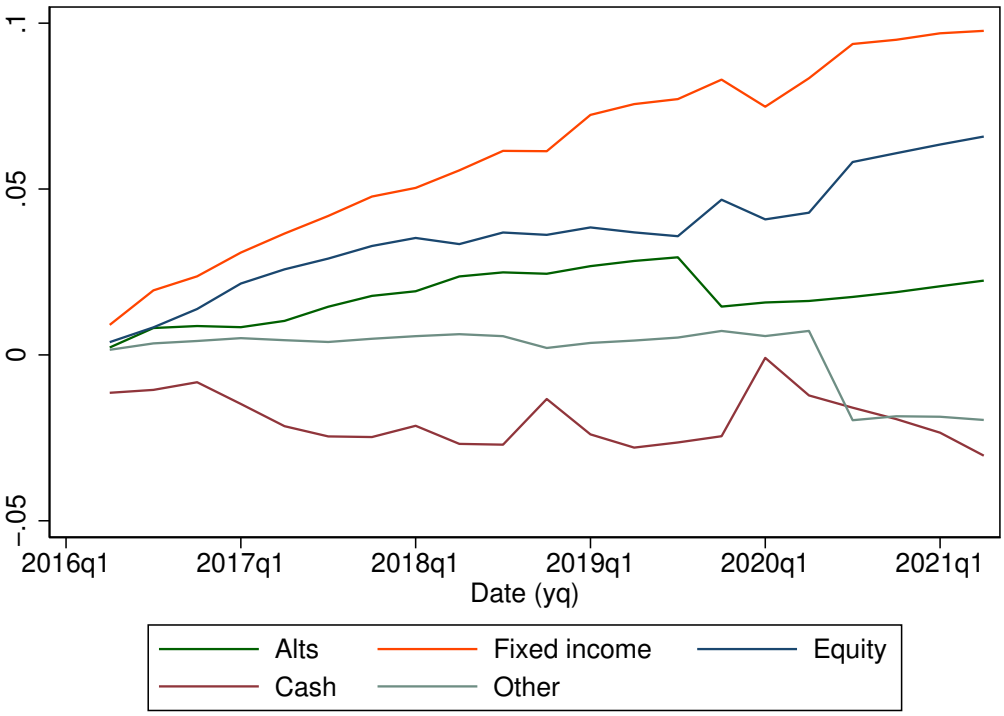


Figure 7: Flows to broad asset classes
We plot the flow into broad asset classes during our sample period from 2016.Q1 to 2021.Q2. Flows are scaled by total assets.



reallocate capital to riskier, higher-yielding assets during the low-rate environment. That said, the average cumulative flows are quite small and amount to about 5-10% of assets over a 5-year period.

3.2 Aggregate flows to liquid risky assets and cash

In terms of flows, our main focus is on the allocation and reallocation of capital to liquid risky asset classes and cash. As a starting point, we explore two aggregate flow measures for each investor in a given quarter. The first measures the flow to cash, which includes the flow to bank accounts and money market mutual funds. We denote this flow by f_{it}^{Cash}. The second measures the aggregate flow to liquid risky asset classes. We denote this flow by $f_{it}^{\text{Liq}} := \sum_{n \in \mathcal{L}} f_{int}$, where $\mathcal{L}$ is the set of liquid risky asset classes as defined in Section 2.2.

In Figure 8, we plot the average value of f_{it}^{Cash} and f_{it}^{Liq} across investors in a given quarter. Three observations stand out from these series. First, the flows to cash and liquid risky assets are strongly negatively correlated: the time-series correlation is -56.5%. This implies that cash is an important substitute for liquid risky assets. Second, the flow to liquid risky assets falls during times of financial market turmoil, such as the last the quarter of 2018 and the first quarter of 2020, while the flow to cash picks up during those same periods. This highlights the role that cash plays as a

Figure 8: Dynamics of the flow to cash and liquid risky assets
We plot the average flow to liquid risky assets, $\frac{1}{I}\sum_i f_{it}^{\text{Liq}}$, in green and the average flow to cash, $\frac{1}{I}\sum_i f_{it}^{\text{Cash}}$, in orange. The sample period is from 2016.Q1 to 2021.Q2.

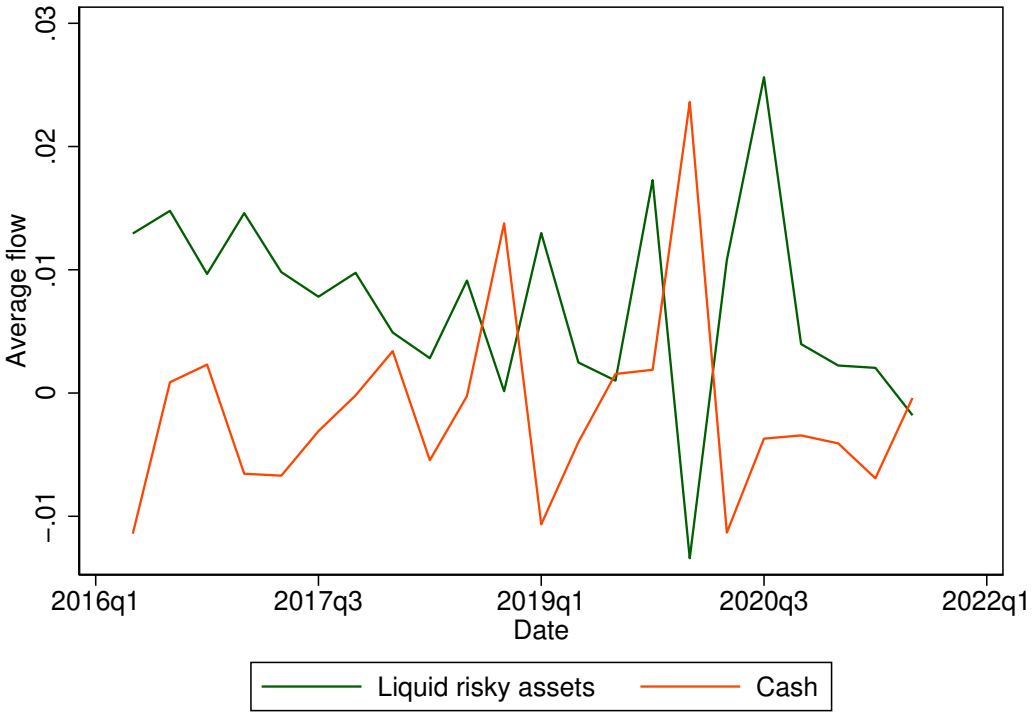
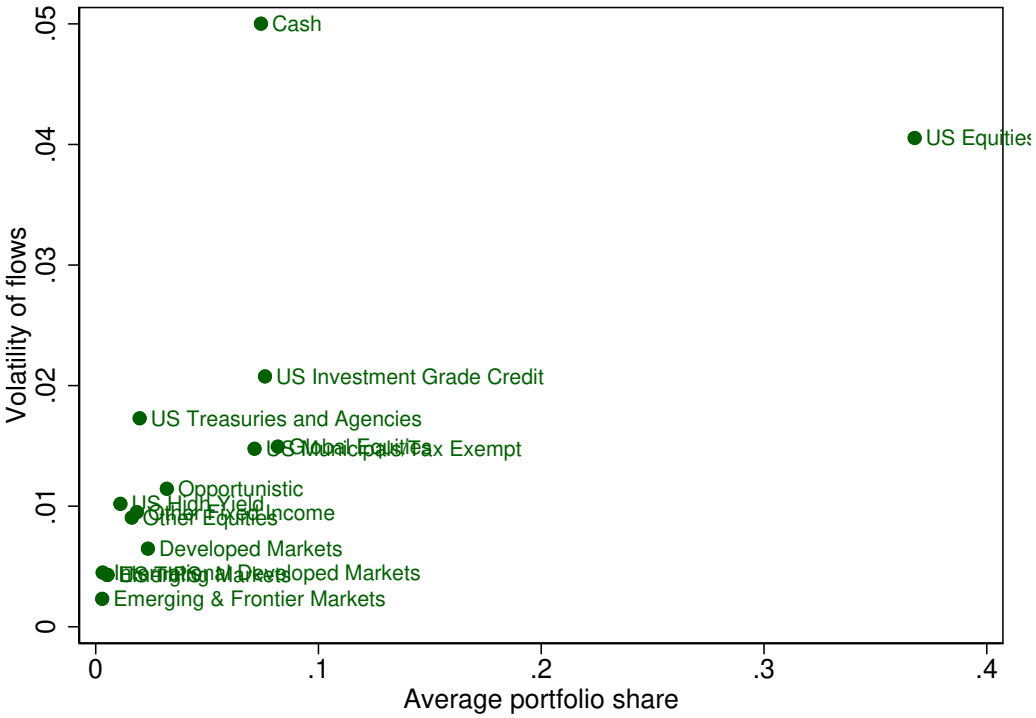


Figure 9: Portfolio shares and the volatility of flows
We plot the average portfolio share allocated to liquid risky asset classes and cash on the horizontal axis and the volatility of flows to the same asset classes on the vertical axis. The sample period is from 2016.Q1 to 2021.Q2.

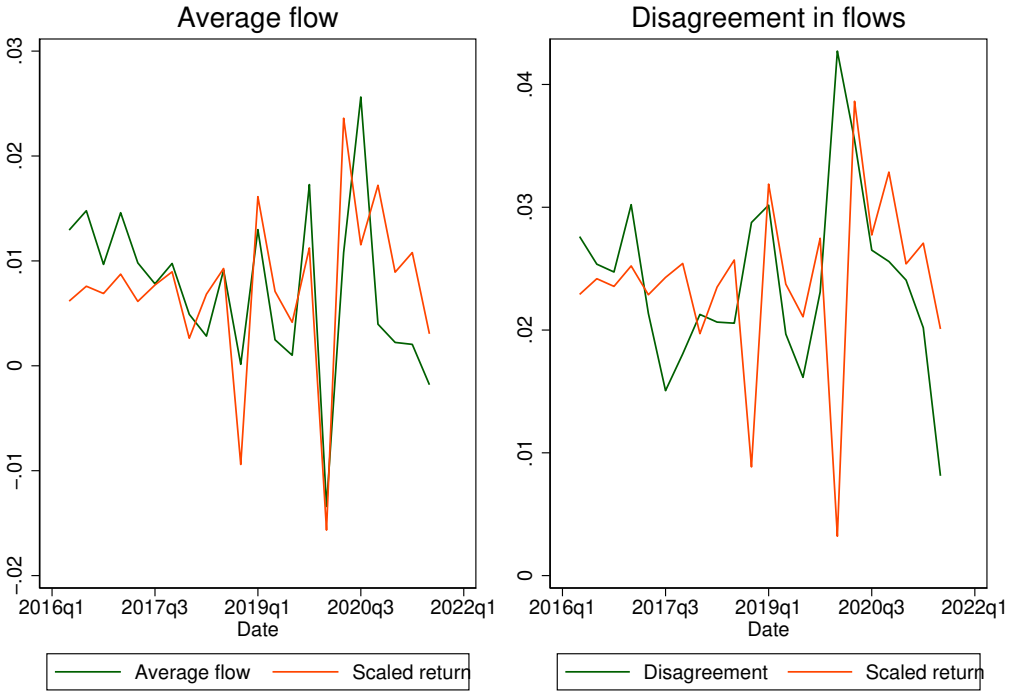


safe asset in investors' portfolios.

Third, the flow to cash is about as volatile as the flow to all liquid risky assets. Both series have a quarterly volatility of 0.8%. Yet, the average cash share is only 7.4% versus 74.1% for the fraction invested in liquid risky assets. This implies that cash is disproportionally volatile. As another way to illustrate the volatility of flows to cash, we plot the average share invested in a particular asset class on the horizontal axis and the (quarterly) standard deviation of flows on the vertical axis in Figure 9. We measure both moments across all investors and quarters. For all asset classes except cash, the volatility of flows aligns closely with the average fraction invested in that asset class; the quarterly volatility of flows is about 10% of the average fraction invested in the asset class. Using this simple metric, we would expect the flow to cash to be less than 1% per quarter, but we find it to be close to 5%.

Economically, the reason is that cash serves two purposes. First, as we discussed before, cash serves as a safe asset: the flows are strongly negatively correlated with the flows to liquid risky assets and the flow to cash increases during times of stress while the opposite is true for the flow to liquid risky assets. If this was the dominant force explaining fluctuations in the flow to cash, however, we would not expect to see this excessively volatile flow.

Figure 10: Flows to liquid risky assets, returns, and disagreement
In the left panel, we plot the time series of f_{it}^{Liq}, averaged across investors in a given quarter, alongside the return on the US stock market from CRSP. We adjust the mean and standard deviation of the return series to match those of the flow series. In the right panel, we plot the disagreement in flows to liquid risky asset classes, as measured by the inter-quartile range of f_{it}^{Liq} across investors in a given quarter, alongside the return on the US stock market. As before, we adjust the mean and standard deviation of the return series to match those of the disagreement series.

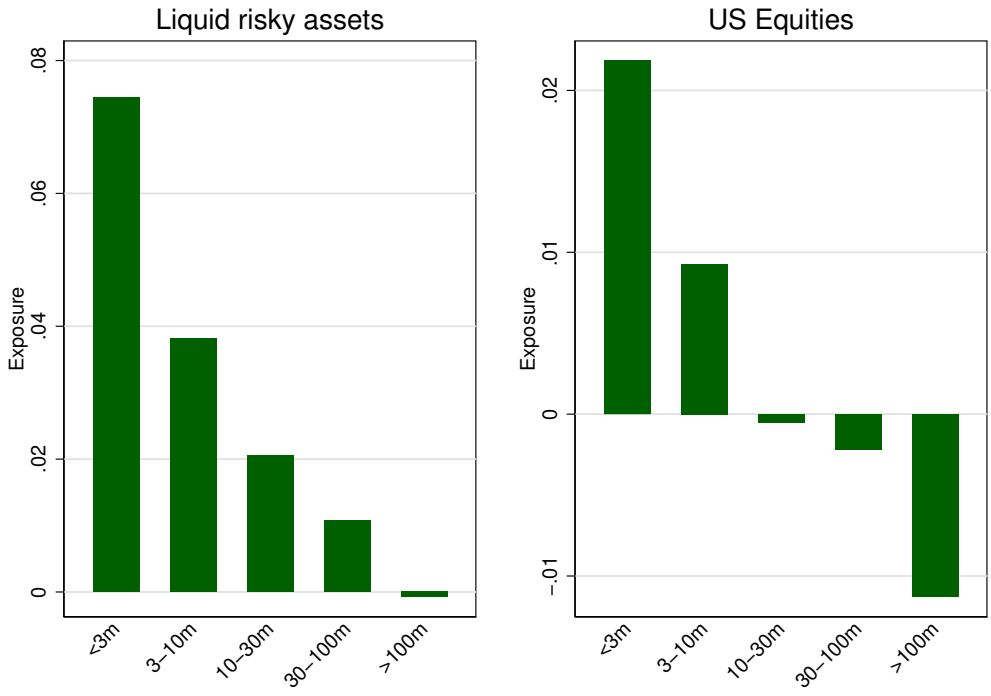


In addition to serving the role of a safe asset, cash holdings are used to buffer liquidity shocks as well. Those volatile liquidity shocks affect the flow to cash but they do not affect the flow to liquid risky assets. This separate determinant of flows to cash adds volatility and this determinant is empirically important. Given this dual role that flows to cash play, we analyze these flows separately from f_{it}^{Liq}.

Next, we explore the link between market conditions and the flow to liquid risky assets in more detail. We first plot the time series of f_{it}^{Liq}, again averaged across investors in a given quarter, alongside the return on the aggreate US stock market from CRSP in the left panel of Figure 10. We adjust the mean and standard deviation of the return series to match those of the flow series. The two series are strongly positively correlated; the time-series correlation between the average flow and US stock market returns is 62%.

In the right panel of Figure 10, we plot the disagreement across investors as measured by the inter-quartile range of f_{it}^{Liq} across investors in a given quarter. We also plot this series alongside the

Figure 11: Exposure of flows to aggregate returns by wealth group
In the left panel, we plot the slope coefficients of a regression of flows to liquid risky assets on the aggregate return on the US stock market by wealth group. In the right panel, we plot the slope coefficients of a regression of flows to US equities on the aggregate return on the US stock market by wealth group. The sample period is from 2016.Q1 to 2021.Q2.



return on the US stock market, adjusting the mean and standard deviation of the return series to match those of the disagreement series as before. In this case, we find that the correlation is -13.3%, implying that disagreement goes up during market downturns. This pattern is particularly salient on the downside during the two most extreme quarters in our sample, that is, the last quarter of 2018 and the first quarter of 2020.

Motivated by the correlations between the average flow, disagreement in flows, and the return on the US stock market, we conclude this section by analyzing how flows connect to stock market returns across the wealth distribution. To this end, we first compute the average of f_{it}^{Liq} across investors in a given wealth group and quarter. We then regress the average flow for this wealth group on the US stock market return.

In the left panel of Figure 11, we plot the estimated slope coefficient for each of the wealth groups. Remarkably, we find that the slopes monotonically decline in wealth. In fact, for the households with more than $100 million in assets, the slope turns slightly negative. This implies that less wealthy households sell during market downturns and thus act pro-cyclically. This behavior amplifies stock market fluctuations.

For the wealthy households, the interpretation is less clear cut. One potential interpretation is that wealthy households are more inert and do not respond at all to turmoil in the markets. Such inelastic behavior would indirectly contribute to amplifying demand shocks of other investors by lowering the elasticity of the overall market (Gabaix and Koijen, 2021). Another interpretation is that wealthy investors instead provide elasticity to the stock market and reallocate capital from fixed income asset classes to equities, thereby leaving the overall flow to liquid risky asset classes largely insensitive to market returns.

To separate these hypotheses, we zoom in on US equities, which is the largest liquid risky asset class and the asset class that best captures how investors respond to fluctuations in the US stock market. We repeat the same analysis as before, but now using f_{int} for US equities instead of f_{it}^{Liq}.

In the right panel of Figure 11, we plot the estimated slope coefficients for each of the wealth groups. As in the right panel, the less wealthy households (those with assets below \$10 million) act pro-cyclically. However, the new insight that emerges from this figure is that (ultra) high net worth households provide elasticity to the market by buying equities during economic downturns. This rejects the idea that wealthy households are fully inert and supports the notion that they actively reallocate capital.

Despite the striking pattern, we note that the empirical magnitude of the effect is quite modest. A 10% decline in the stock market leads to a 0.1% inflow into equities for very rich households (over \$100 million in assets), and a -0.2% inflow (i.e., an outflow) for relatively poorer (less than \$3 million in assets). So while the main qualitative takeaway is that wealthy households provide elasticity to the market, the main quantitative takeaway is that the magnitudes are small.

3.3 Decomposing flows

The analysis in the previous section focused largely on the flow to cash and all liquid risky assets. In this section, we extend the analysis by studying how households allocate capital across liquid risky asset classes.

To this end, we develop a simple framework that allows us to use principal components analysis (PCA) to measure how investors reallocate capital from one asset class to another. As a starting point, we remove the factors that we analyzed in the previous section via the following regression

$$f_{int} = \alpha_n + \beta_n f_{it}^{\text{Liq}} + \gamma_n f_{it}^{\text{Cash}} + f_{int}^{\perp}. \tag{4}$$

Using the fact that $f_{it}^{\text{Liq}} = \sum_{n\in\mathcal{L}} f_{int}$, it holds that $\sum_{n\in\mathcal{L}} \beta_n = 1$ and thus $\sum_{n\in\mathcal{L}} \alpha_n = \sum_{n\in\mathcal{L}} \gamma_n = \sum_{n\in\mathcal{L}} f_{int}^{\perp} = 0$. In this regression, we are primarily interested in the residuals, $f_{int}^{\perp}$. The property that $f_{int}^{\perp}$ sum to zero across all liquid risky asset classes makes it an appealing measure of rebalancing flows. Indeed, if $f_{it}^{\text{Liq}} = f_{it}^{\text{Cash}} = 0$, then all rebalancing across asset classes is captured by $f_{int}^{\perp}$.

The coefficients also have a natural interpretation. The slope on f_{it}^{Liq}, β_n, measures how fresh

flows to liquid risky assets are allocated to asset class n. If households maintain fairly stable portfolio shares over time, we expect $\beta_n \simeq \mathbb{E}[\theta_{int}]$, that is, capital is allocated in proportion to existing portfolio shares. The slope on f_{it}^{Cash}, γ_n, measures how flows to cash may be correlated to flows to a particular asset class. For instance, if Treasuries are a close substitute to money market funds during times of stress, as they are both safe assets albeit with different durations, we expect γ_n to be negative. The intercept, α_n, measures broad reallocation trends during our sample period. Empirically, both α_n and γ_n are economically small and we will not explore them in further detail in the remainder of this section.

In the second step of the analysis, we model the rebalancing flows, $f_{int}^{\perp}$, using a factor model

$$f_{int}^{\perp} = \sum_{k} \lambda_{it}^{(k)} \eta_n^{(k)} + u_{int}, \tag{5}$$

where $k = 1, ..., K$ indexes the number of factors. We estimate the factor model using principal components analysis (PCA). Economically, as $\sum_{n \in \mathcal{L}} \eta_n^{(k)} = 0$, these coefficients represent a long-short trading strategy, for instance, purchase US equities and sell Treasuries. We are therefore particularly interested in measuring $\eta_n^{(k)}$ as they summarize the key rebalancing dimensions in the data. The loadings, $\lambda_{it}^{(k)}$, capture the exposure of investor i in quarter t to factor k. The residual, u_{int}, capture the idiosyncratic rebalancing decisions of an investor.

We first report the estimates of β_n alongside the average portfolio shares in Figure 12 for each of the liquid risky asset classes. As discussed before, if investors maintain fairly constant shares, we expect $\beta_n \simeq \mathbb{E}[\theta_{int}]$. The figure shows that the estimates of β_n are closely aligned with $\mathbb{E}[\theta_{int}]$. This implies that, at least on average, constant portfolio shares is a reasonable way to model demand.

In the next step, we estimate the factor model. In Figure 13, we summarize the fraction of the variance in $f_{int}^{\perp}$ explained by various factors. As the figure makes clear, there are important common components and the first three factors explain about 65% of the variation in portfolio rebalancing. We now explore the properties of those rebalancing factors.

In Figure 14, we report the estimates of $\eta_n^{(k)}$ for the first three factors, which explain about 65% of the variation in $f_{int}^{\perp}$. As we discussed before, these loadings have the convenient property that $\sum_{n \in \mathcal{L}} \eta_n^{(k)} = 0$, which means that they can be interpreted as long-short portfolios or trades.

The factors have a clear economic interpretation. The first factor rebalances from US equities to long-duration fixed income, such as US investment-grade corporate bonds, Treasuries and agencies, and municipal bonds. This factor therefore captures the long-term equity risk premium.

The second factor rebalances from US investment grade bonds to US Treasuries and agencies. This factor therefore captures the credit spread. The third factor is a combination of two economically interpretable trades. The first rebalances from US Treasuries and agencies and US investment-grade corporate bonds to US municipal and tax exempt bonds. We have seen before that municipal bonds play a nontrivial role in households' portfolios, in particular for wealthier

Figure 12: Allocation of new flows
We plot the estimates of β_n in equation (4) for all the liquid risky asset classes. We compare the estimates to the average portfolio shares, $\mathbb{E}[\theta_{int}]$. If investors maintain stable shares invested in the different asset classes, then we expect $\beta_n \simeq \mathbb{E}[\theta_{int}]$. The sample period is from 2016.Q1 to 2021.Q2.

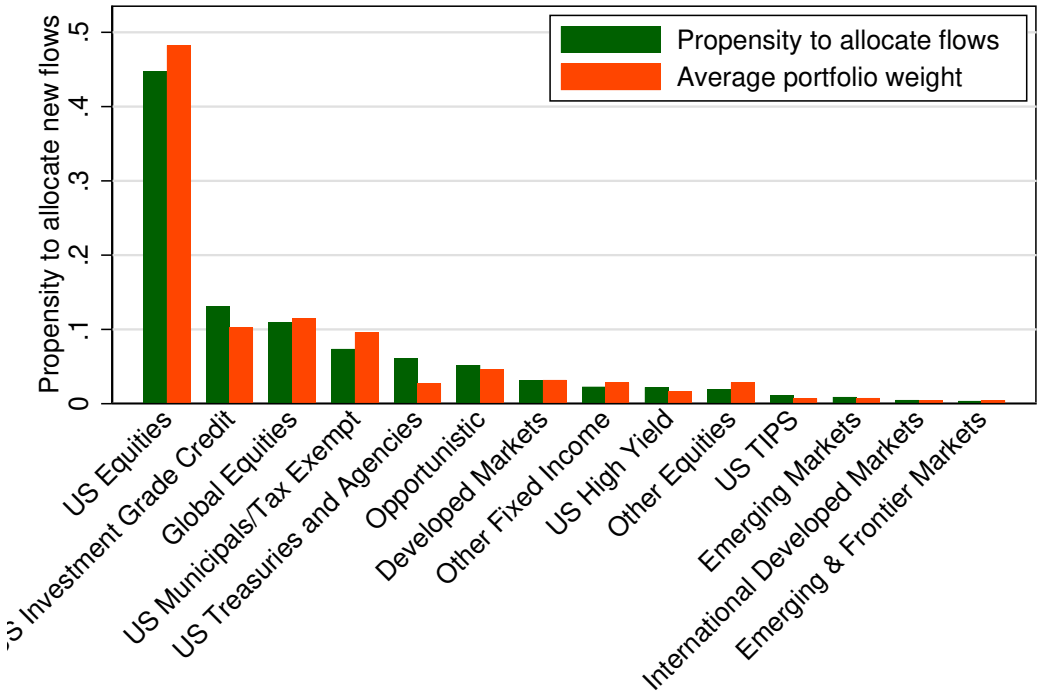


Figure 13: Factor structure in portfolio rebalancing
We plot the share of variance of $f_{int}^{\perp}$ explained by the principal components, see equation (5). The sample period is from 2016.Q1 to 2021.Q2.

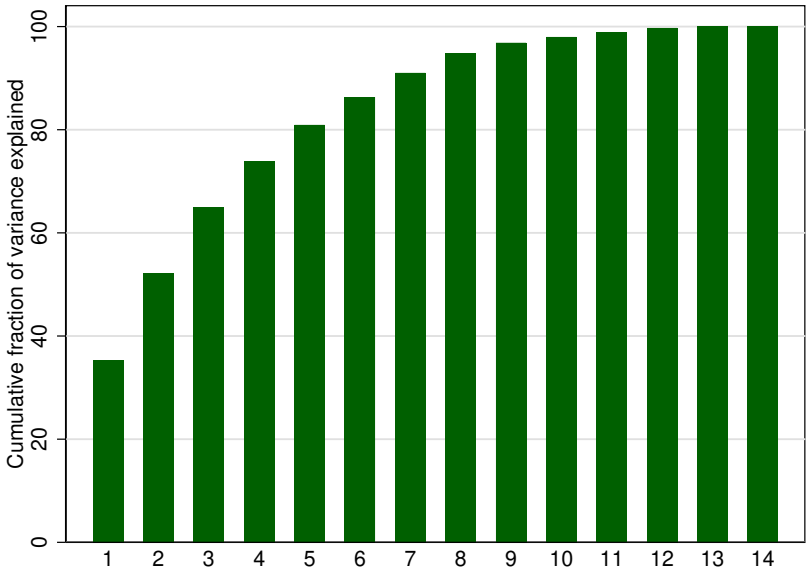
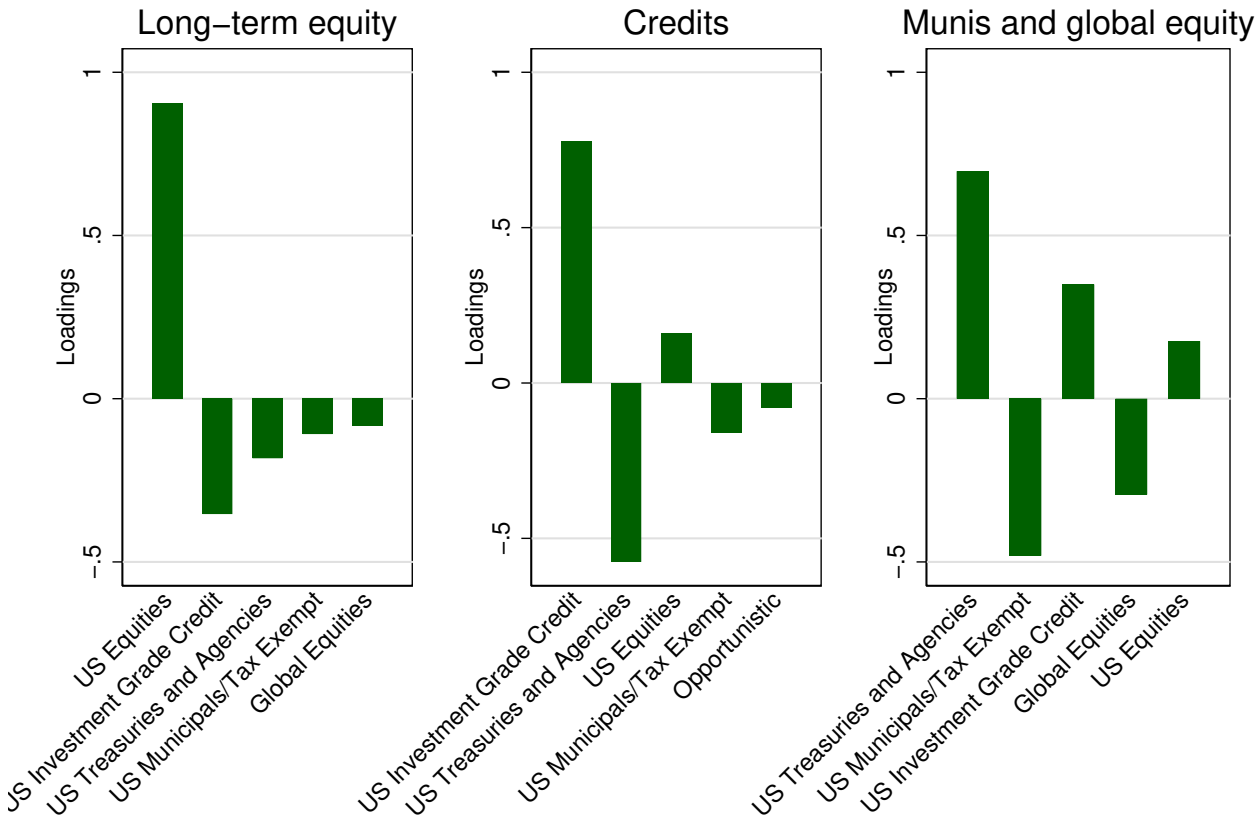


Figure 14: Rebalancing exposures across asset classes
We report the estimates of $\eta_n^{(k)}$, for $k = 1, 2, 3$, in equation (5). The coefficients capture the main rebalancing directions based on $f_{int}^{\perp}$. The sample period is from 2016.Q1 to 2021.Q2.



investors. The second leg of this factor rebalances from global equities to US equities. Hence, the third factor captures a combination of the risk premium in municipal debt markets relative to other safe fixed income markets and the risk premium on global equities relative to US equities.

Taken together, our decomposition of flows into investor-level factors, f_{it}^{Cash} and f_{it}^{Liq}, and rebalancing factors, $f_{int}^{\perp}$, provides qualitative and quantitative guidance to design of macro-finance models with rich heterogeneity.

4 Asset demand across US equities

We have so far focused on rebalancing across asset classes. In this section, we focus in more detail on the largest liquid risky asset class: US equities. In Section 4.1, we discuss the construction of the data and we provide summary statistics about the average composition of investors' portfolios within US Equities. We then study how households' active risk taking varies with wealth in Section 4.2. We also compare households' active share to the active share of institutions. In Sections 4 - 4.5, we estimate a demand system for the cross section of US equities.

4.1 Data construction

We summarize how we construct our sample of US equity holdings for households. We start from a merged dataset of equity prices from CRSP, firm fundamentals from Compustat, and institutional holdings as constructed in Koijen et al. (2021).[11] The institutional holdings data are extracted from SEC form 13F filings. We also use the Koijen and Yogo (2019) instrument, which we discuss in more detail below, for the household sector in these data. We refer to Koijen et al. (2021) for details on the data construction.

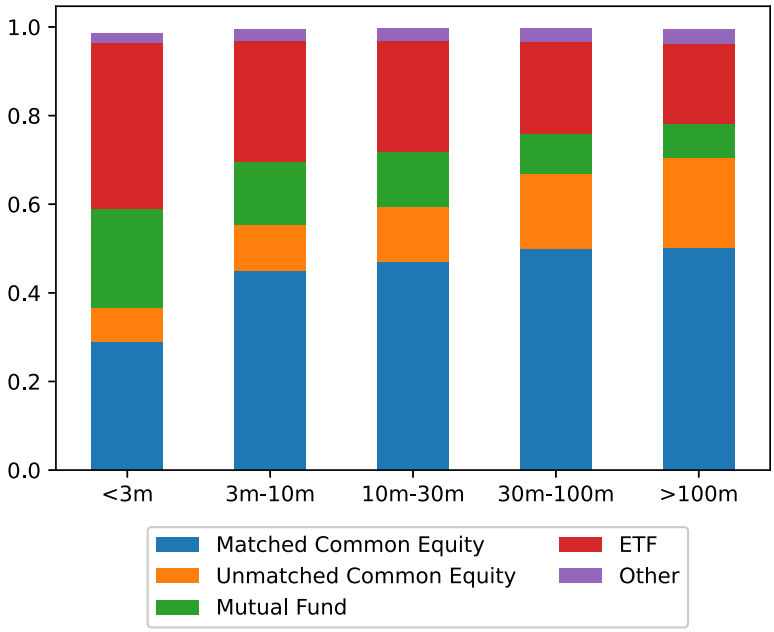
For each quarter of our sample, we merge these data on CUSIP with the security-level holdings data from Addepar.[12] While a security's CUSIP may change over time, the CUSIP in the Addepar data is fixed to its value when the security appears in the portfolio of an investor. To match the data, we compile a list of CUSIPs that are ever associated with a security's permanent identifier, which is PERMNO in CRSP and FSYM_ID in FactSet. We then match the Addepar data to the merged CRSP-Compustat-FactSet data.

We summarize the match quality in Figure 15. There are five possible categories: (i) equities that we can match with the Koijen et al. (2021) data ("Matched Common Equity"), (ii) equities that we cannot match yet that are classified as Common Equity in Addepar ("Unmatched Common

[11] The data construction closely follows Koijen and Yogo (2019), but Koijen et al. (2021) use holdings data from FactSet instead of Thomson Reuters.

[12] As in Section 2.2, when holdings are classified as "Unknown Equities" and the investment sub type is either a mutual fund or ETF, we use the Morningstar classification to assign the holdings to "US Equities" if the Morningstar classification is "U.S. Equity."

Figure 15: Average portfolio weight
We plot the decomposition of US equities by wealth group in 2019.Q4. The "Matched Common Equity" share corresponds to the holdings that can be merged with the CRSP-Compustat-FactSet data. The "Unmatched Common Equity" share corresponds to securities that are classified by Addepar as Common Equity but that cannot be matched to the CRSP-Compustat-FactSet data. The "Mutual Fund" and "ETF" shares correspond to US equity positions held via those funds. The remaining positions are classified as "Other."

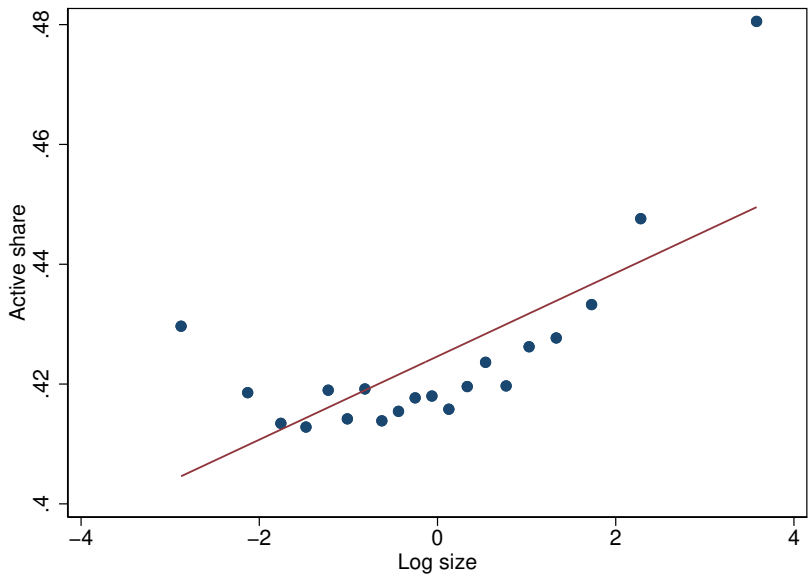


Equity"), (iii) equity holdings via mutual funds, (iv) equity holdings via ETFs, and (v) a residual category (e.g., preferred equity). We report in Figure 15 the average portfolio weight of investors within US equities and by wealth group.

There are two main takeaways. First, except for households with wealth under $3 million, direct equity holdings is the largest category. Mutual funds and ETFs are large as well. In a future version of the paper, we explore the impact of "looking through" the holdings of mutual funds and ETFs. Second, the average portfolio share in directly-held common stocks increases from around 40% for investors with wealth below $3m to more than 70% for investors with wealth above $100m. This increase is essentially offset by a decrease in indirect positions held through mutual funds and ETFs. In other words, (ultra-)high-net-worth individuals rely less on widely-available forms of delegated portfolio management.

To arrive at our final data set, we make a series of adjustments to correct for infrequent data errors and to ensure that we focus on investors with a broad enough investment universe. We proceed in four steps. First, we retain portfolios with wealth across all asset classes greater than or equal to $100,000. Second, we exclude strictly negative positions. Third, we move stocks that are held by only 50 investors to the outside asset to avoid that we focus on small and rarely held

Figure 16: Relation between Active Share and log wealth
We plot households' log wealth (in millions) on the horizontal axis and the corresponding active share on the vertical axis. The results are presented for 2019.Q4.



stocks. Fourth, we restrict attention to investors with at least 20 strictly positive positions.

4.2 Active risk taking across the wealth distribution

To lead off our analysis, we explore how active risk taking varies across the wealth distribution by computing the active share for each investor as

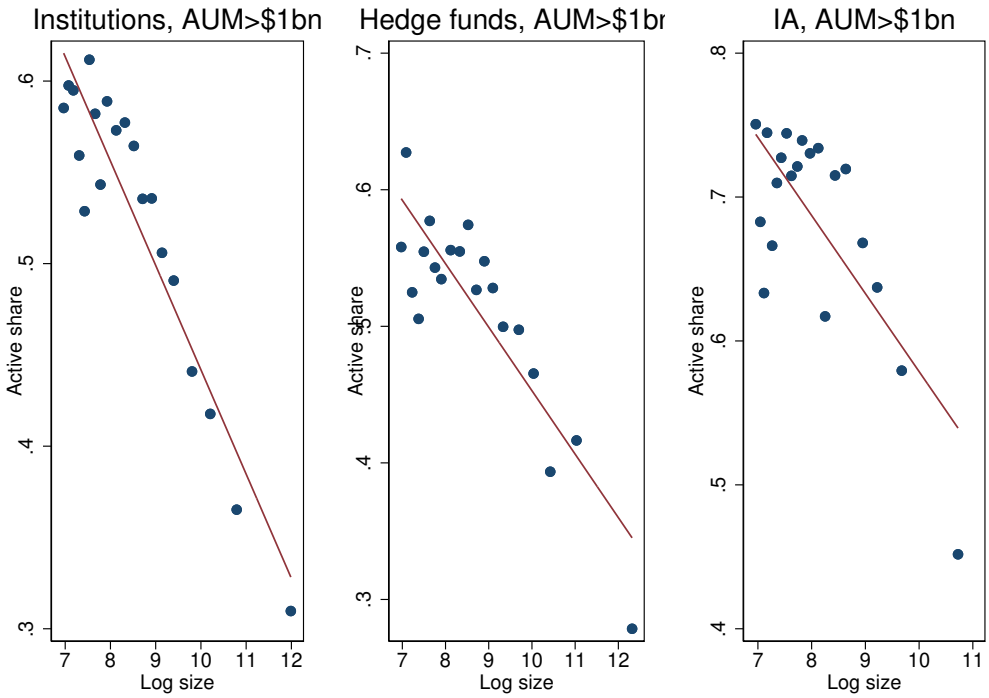
$$AS_{it} = \frac{1}{2} \sum_{a} \left| \theta_{iat}^{\text{US Eq}} - \theta_{iat}^{m} \right|,$$

where $\theta_{iat}^{\text{US Eq}} = \frac{A_{iat}}{A_{it}^{\text{US Eq}}}$ denotes the weight on stock a in the portfolio of US equities for investor i, $A_{it}^{\text{US Eq}} = \sum_{a \in \text{US Eq}} A_{iat}$, and θ_{iat}^{m} denotes the weight on the market portfolio based on the stocks held by investor i at time t. In Figure 16, we compare AS_{it} to households' total wealth using a binned scatter plot. The main takeaway is that wealthier investors choose more active portfolios.

This household-level evidence contrasts the evidence for institutional investors, for which the opposite relation holds and for which the active share decreases in assets under management (Cremers and Petajisto, 2009). We confirm this result in our sample, and in particular if the smallest institutions are excluded (those with assets under management under $1bn).

In Figure 17, we plot the relationship between size and active share for all 13F institutions in the left panel; for hedge funds in the middle panel; and for investment advisors in the right panel. In all cases, the relationship between size and active share is negative. The lower active share for

Figure 17: Relation between Active Share and log size of institutions
We plot investors' log assets (in millions) on the horizontal axis and the corresponding active share on the vertical axis. The results are presented for 2019.Q4.



larger institutions can be explained by the fact that those institutions offer a larger share of passive, index-like products. More generally, it is more challenging to implement active, dynamic strategies for larger asset managers.

4.3 A characteristics-based asset demand system

We now discuss the empirical specification of the asset demand curve that we will estimate. This specification follows Koijen and Yogo (2019), who also provide a micro foundation that we summarize briefly. We start from the assumption that investors have mean-variance preferences with potentially heterogeneous beliefs about expected returns and risk. The demand curve is then given by

$$\theta_{it}^{\text{US Eq}} = \frac{1}{\gamma_i} \Sigma_{it}^{-1} \mu_{it},$$

where Σ_{it} denotes the conditional covariance matrix of returns, μ_{it} the vector of conditional expected returns, and γ_i the coefficient of relative risk aversion.

This model is quite intractable to directly take to the data. However, under two commonly-made assumptions in asset pricing, the demand curve becomes tractable. First, we assume that returns follow a factor model. Second, the factor loadings and the expected returns of a particular stock depend only on the characteristics of that stock. Under these assumptions, the demand for stock a depends only on the characteristics of stock a and a common scalar. This common scalar encodes all information in terms of prices and characteristics of all other stocks.

With this potential micro-foundation in mind, we model the demand of investor i for stock a at time t as

$$\theta_{iat}^{\text{US Eq}} = \frac{\delta_{iat}}{1 + \sum_h \delta_{iht}},$$

where

$$\ln \delta_{iat} = \beta_{0it} me_{at} + \beta_{1it}' x_{at} + \epsilon_{iat}, \tag{6}$$

and me_{at} is the log market capitalization of stock a in quarter t, x_{at} a vector of stock characteristics (including a constant), and ϵ_{iat} latent (unexplained) demand. In our demand specification below, the vector x_{at} includes a stock's market beta, log book equity, log asset growth (or investments), operating profits to book equity, and annual dividends to book equity. These characteristics are motivated by the Fama and French (2015) five-factor model and capture important variation in risk and expected returns.

Because $\sum_a \theta_{iat}^{\text{US Eq}} = 1$, the portfolio weight invested in the outside asset satisfies

$$\theta_{i0t}^{\text{US Eq}} = \frac{1}{1 + \sum_h \delta_{iht}},$$

implying $\ln \delta_{iat} = \ln \frac{\theta_{iat}^{\text{US Eq}}}{\theta_{i0t}^{\text{US Eq}}}$. For the construction of the outside asset, we include all assets in US equities that are not matched to the merged FactSet-CRSP-Compustat data. This implies that the outside asset will include, for instance, any position in mutual funds and ETFs. To ensure that we correctly measure the outside asset, we exclude investors with outside assets under management that account for at least 5% of total assets under management.

4.4 Empirical strategy

The model implies the following estimating equation

$$\ln \frac{\theta_{iat}^{\text{US Eq}}}{\theta_{i0t}^{\text{US Eq}}} = \beta_{0it} me_{at} + \beta_{1it}' x_{at} + \epsilon_{iat}. \tag{7}$$

We assume throughout that characteristics are exogenous to latent demand, $\mathbb{E}[\epsilon_{iat} \mid x_{at}] = 0$. However, we cannot assume that latent demand is uncorrelated with prices, as high latent demand leads to higher prices via market clearing.

We use the instrument as proposed by Koijen and Yogo (2019). In particular, they assume that there is an exogenous component to investors' choice sets, which can be driven by investment mandates, investors' expertise, et cetera. The instrument is then constructed as

$$me_{iat} = \ln \left(\sum_{i \neq j, i \neq HH} A_{it} \frac{\mathbb{I}_{jat}}{1 + \sum_h \mathbb{I}_{hat}} \right),$$

where A_{it} are the assets of institution i and $\mathbb{I}_{jat}$ equals one if stock a is in the choice set of investor j and zero otherwise. The choice set of an investor consists of all stocks held during the current and previous 11 quarters. Koijen et al. (2021) show that the estimates are robust to changes in the length of the window; either by extending the window further backwards or even forward.

To estimate the model, we pool investors as many households hold a fairly small number of stocks in their portfolios and we cannot reliably estimate their demand curves in each quarter. We propose a new clustering algorithm to pool investors.

To explain conceptually how we cluster investors, we start from a linearized demand curve of investors (Gabaix and Koijen, 2021),

$$\hat{q}_{ia} = -\gamma_{0i} \hat{p}_a + \beta_{1i}' \hat{x}_a + \hat{\epsilon}_{ia},$$

where we omit time subscripts for expositional simplicity. The variables with a hat, $\hat{x}$, indicate deviations from baseline values, for instance, $Q_{ia}' = Q_{ia}(1 + \hat{q}_{ia})$. In this notation, $\gamma_{0i} \simeq 1 - \beta_{0i}$. Market clearing implies $\hat{q}_{Sa} = 0$, where $\hat{x}_{Sa} = \sum_i S_{ia} \hat{x}_{ia}$ is the size-weighted average of $\hat{x}_{ia}$, and S_{ia}

the fraction of asset a held by investor i. This implies for the market-clearing price

$$\hat{p}_a = \frac{\beta_{1S}' \hat{x}_a + \hat{\epsilon}_{Sa}}{\gamma_{0S}}.$$

Substituting the market-clearing price into the demand equation, we obtain

$$\hat{q}_{ia} = \left(\beta_{1i} - \frac{\gamma_{0i}}{\gamma_{0S}}\beta_{1S}\right)' \hat{x}_a + \hat{\epsilon}_{ia} + \frac{\hat{\epsilon}_{Sa}}{\gamma_{0S}}.$$

This shows that investors with similar demand, at least in terms of observable characteristics, can be identified by regressing demand on characteristics. The slope of this regression identifies $\widetilde{\beta}_{1i} = \beta_{1i} - \frac{\gamma_{0i}}{\gamma_{0S}}\beta_{1S}$. We then cluster households by $\widetilde{\beta}_{1i}$.

Concretely, we regress $\ln \delta_{iat}$ on characteristics and obtain the slopes. We standardize the coefficients and then use an iterative nearest neighbor algorithm to form clusters of investors who jointly hold at least 2,000 stocks. [13] We then pool investors in the same cluster in estimating the demand curves.

4.5 Empirical results

We report in Figure 18 the average price elasticity (for individual stocks), which we approximate by $1 - \beta_{0it}$, by wealth group. We report the results for 2019.Q4, but this pattern holds for all other quarters as well.

The estimated elasticities range from 0.58 to 0.82, implying that a 1% price shock decreases asset demand by 0.58%-0.82%. Moreover, the estimated demand elasticities decline in wealth, implying that wealthier investors rebalance their portfolios less aggressively in response to price shocks. Koijen et al. (2021) document a similar negative relationship between size and demand elasticities for institutional investors.

In Figure 19, we plot the time-series of average demand elasticity across all investors. Individual investors have become more inelastic between 2016Q1 and 2021Q2. Moreover, elasticities decline during the onset of the turmoil surrounding the COVID-19 pandemic, which is consistent with evaporating liquidity during times of stress.

5 The demand elasticity of asset class level flows

We studied the dynamics of flows across various asset classes in Section 3. Flows either reflect a response to prices (that is, changes along the demand curve) or shifts in the demand curve due

[13] We start from an unassigned investor that holds the highest number of stocks at each iteration. We then assign those unassigned investors that are most similar in terms of $\tilde{\beta}_{1it}$ as measured by the Euclidean distance to the same cluster.

Figure 18: Demand elasticity by wealth group
We plot the average demand elasticity, which we approximate by $1 - \beta_{0i}$, by wealth group. The results are presented for 2019.Q4.

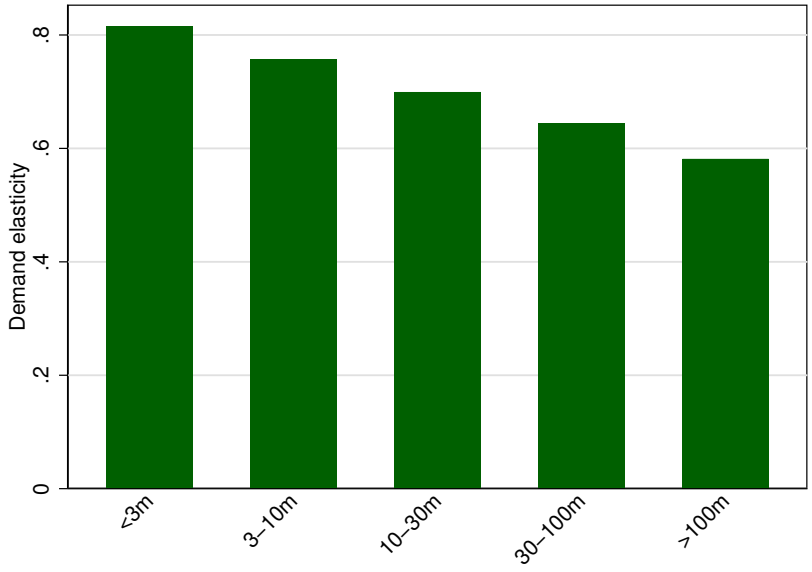
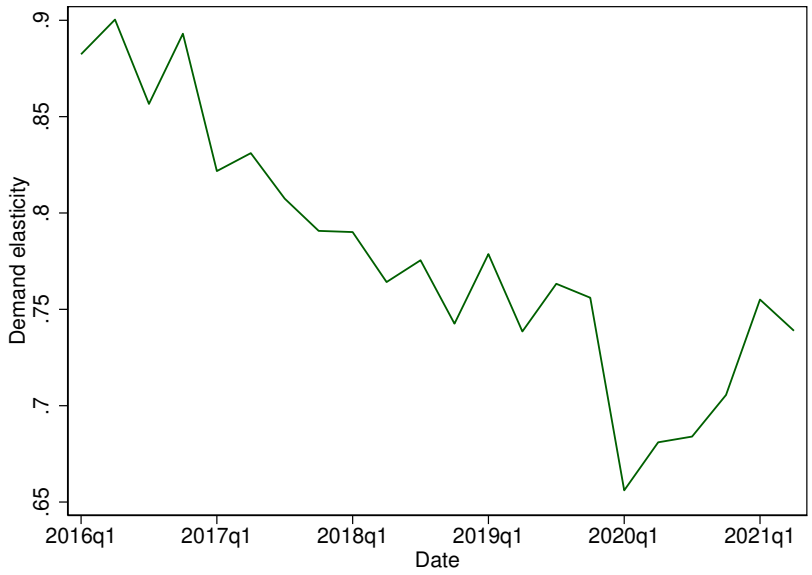


Figure 19: Time series of demand elasticity
We plot the average demand elasticity, which we approximate by $1 - \beta_{0i}$, over time. The sample period is from 2016.Q1 to 2021.Q2.



to, for instance, changes in perceived risk or risk aversion. In this section, we propose a model to separate those two terms. This is a nontrivial task as returns and demand shocks are plausibly correlated, and we provide a new way of estimating asset-class level demand elasticities.

5.1 An asset-class level model of demand

We refer to asset classes as n. We also define $q_{int} = \frac{F_{int}}{A_{in,t-1}+\frac{1}{2}F_{int}}$, which is the asset class level flow using the Davis and Haltiwanger (1992) measure of flows. By definition, $q_{int} \in [-2, 2]$. Analogously, we define the aggregate flow $f_{it} = \frac{\sum_n F_{int}}{\sum_n A_{in,t-1}+\frac{1}{2}F_{int}}$.[14] The model is implemented only for liquid risky asset classes as before.

We start from the general decomposition

$$q_{int} = -\sum_m \zeta_{inm,t-1} r_{imt} + f_{it} + \nu_{int}, \tag{8}$$

where r_{int} is the return of investor i in asset class n at time t, ν_{int} are demand shocks, and ζ_{inmt} is the (cross) elasticity of investor i in asset class n for a price change in asset class m at time t.

As $\mathbb{E}[r_{imt}\nu_{int}]$ may be non-zero, we cannot estimate the model in (8) by OLS. There are generally two ways to proceed. First, in some case, it may be possible to find instruments for returns and estimate the elasticities using instrumental variables (Koijen and Yogo (2019); Gabaix and Koijen (2020)). Second, we can impose additional structure to identify the elasticities from changes in the structure of the covariance matrix (Rigobon (2003)). We proceed using this second approach, which is new in the literature on asset demand estimation, and motivate the restrictions by economic theory.

To derive the economic restrictions that we impose, we start from a model of portfolio weights

$$\theta_{int} = \frac{P_{int}Q_{int}}{A_{it}} = \frac{\theta_{in,t-1}e^{V_{int}}}{\sum_m \theta_{im,t-1}e^{V_{imt}}},$$

where V_{int} are demand shocks and θ_{int} portfolio weights that sum to one, $\sum_n \theta_{int} = 1$. We approximate $Q_{int} \simeq Q_{in,t-1}(1 + q_{int})$, $P_{int} \simeq P_{in,t-1}(1 + r_{int})$, and $A_{it} = A_{i,t-1}(1 + r_{it} + f_{it})$ where $r_{it} = \sum_n \theta_{in,t-1} r_{int}$. This model implies

$$\begin{aligned} q_{int} &= r_{it} - r_{int} + V_{int} - \sum_m \theta_{im,t-1} V_{imt} + f_{it} \\ &= -\left(\sum_m (1_{n=m} - \theta_{im,t-1})\, r_{imt}\right) + \left(V_{int} - \sum_m \theta_{im,t-1} V_{imt}\right) + f_{it}. \end{aligned}$$

If the shifters V_{int} are 0, then

$$\zeta_{inm,t-1} = 1_{n=m} - \theta_{im,t-1}. \tag{9}$$

[14] We assume throughout that the aggregate flow is not price elastic.

Those are the own- and cross-price elasticities coming from a model where the agent solely wishes to keep proportions across asset classes, and the rebalancing is exclusively due to differences in relative returns.

Next we discuss how to model V_{int}. We seek guidance from the traditional model, with uncorrelated returns, so[15]

$$\theta_{in,t-1}e^{V_{int}^{\text{trad}}} = K\frac{\pi_{nt}}{\sigma_{nt}^2},$$

where we can allow for heterogeneity in beliefs about risk premia, π_{nt}, and perceptions of risk, σ_{nt}. We then have

$$V_{int}^{\text{trad}} + \Delta \ln \sigma_{nt}^2 = \Delta \ln \pi_{nt} = \frac{\Delta \pi_{nt}}{\pi_{n,t-1}} = -\frac{\delta_n + \phi_n}{\bar{\pi}_n} r_{int} = -\psi_n r_{int}, \tag{10}$$

where we use

$$\psi_n = -\frac{d \ln \pi_{int}}{dp_{int}}, \tag{11}$$

and that in a model where disturbances mean-revert at a (continuous-time) rate ϕ_n, we have

$$\psi_n = \frac{\delta_n + \phi_n}{\bar{\pi}_n}, \tag{12}$$

where δ_n is the dividend-price ratio of asset class n and $\bar{\pi}_n$ the long-run expected excess return.

One useful special case to consider can be where $\bar{\pi}_n = \lambda \sigma_n$, where the Sharpe ratio, λ, is the same across asset classes (on average). This specification implies that elasticities increase with volatility, which seems consistent with the data. Also, we note that empirically, the expression (11) is probably easiest to estimate from simple predictive regressions that include price-based predictors.

Putting everything together, we have

$$q_{int} - f_{it} = -\sum_m \zeta_{inm,t-1} r_{imt} + \nu_{int}, \tag{13}$$

where $\nu_{int} \equiv -\left(\Delta \ln \sigma_{nt}^2 - \sum_m \theta_{im,t-1} \Delta \ln \sigma_{mt}^2\right)$, and

$$\zeta_{inm,t-1} = \left(1_{n=m} - \theta_{im,t-1}\right)\left(1 + \psi_m\right). \tag{14}$$

5.2 Identifying assumptions

We make two types of identifying assumptions to estimate (8): one related to demand elasticities and one related to demand shocks.

[15] We can easily extend this approach using a factor model and keep the same tractability.

Demand elasticities We build on the structure in (14), but we allow for additional flexibility in case households do not react as strongly (or more strongly) to prices as predicted by the traditional mean-variance model:

$$\zeta_{inm,t-1} = (1_{n=m} - \theta_{im,t-1}) (\xi_i + \chi_i \psi_m) . \tag{15}$$

This modification reduces the number of parameters to two, (ξ_i, χ_i), and potentially N parameters that are common across investors, $\psi = (\psi_1, ..., \psi_N)$.

The two parameters, ξ_i and χ_i have intuitive interpretations. The parameter ξ_i captures the speed at which households rebalance back to fixed shares in response to relative returns.[16]

The model can be extended to allow for heterogeneous χ_{in} across asset classes. For instance, suppose that for a given movement in equity prices, the investment in corporate bonds is less reactive than that in government bonds, perhaps due to the lower liquidity of corporate bonds. We can capture this pattern by $\chi_{i,\text{Corp Bonds}} < \chi_{i,\text{Gov Bonds}}$. This in turn suggest parametrizing χ_{in} by characteristics such as liquidity or inverse volatility.

Demand shocks We assume the following model for demand shocks

$$\nu_{int} = \sum_m \mu^0_{inm} r_{mt} + \nu^{\perp}_{int}, \tag{18}$$

for some coefficients μ^0_{inm}, and with $\nu^{\perp}_{int}$ uncorrelated with r_{nt}. The key identifying assumptions are that μ^0_{inm} are constant over time and that $\mathbb{E}[\nu^{\perp}_{int} r_{mt}] = 0$ for all n and m.[17] We discuss extensions of this model in Section 5.4.

Empirical specification We put the model together by substituting (15) and (18) in (13), implying

$$\begin{aligned} q_{int} - f_{it} &= -\sum_m \zeta_{inm,t-1} r_{imt} + \nu_{int} \\ &= \sum_m \mu_{inm} r_{imt} + \xi_i \left(\sum_m \theta_{im,t-1} r_{imt} \right) + \chi_i \left(\sum_m \theta_{im,t-1} \psi_m r_{imt} \right) + \nu^{\perp}_{int}, \end{aligned} \tag{19}$$

[16] In a model with slow-moving adjustment where the speed of adjustment is ϕ_i, we have:

$$\xi_i = \phi_i, \qquad \chi_i = \phi_i \kappa_i, \tag{16}$$

where κ_i measures the response to risk premia, while the long-run elasticities are:

$$\zeta^{LR}_{inm,t-1} = (1_{n=m} - \theta_{im,t-1}) (1 + \kappa_i \psi_m) . \tag{17}$$

It is the long run elasticities that matter for much of asset pricing, at least when markets are not too myopic.

[17] The constant loadings can be viewed as a version of Rigobon (2003)'s idea to identify elasticities via heteroskedasticity. Constant loadings impose constraints on the covariance matrix, and fluctuations in θ_{int} can now be used for identification.

where $\mu_{inm} \equiv \mu^0_{inm} - 1_{n=m}(\xi_i + \chi_i \psi_m)$ are coefficients that we do not seek to structurally interpret.

5.3 Econometric approach

The model in (19) leads us to define

$$r^\theta_{it} \coloneqq \sum_m \theta_{im,t-1} r_{imt}, \qquad r^\psi_{it} \coloneqq \sum_m \theta_{im,t-1} \psi_m r_{imt}, \tag{20}$$

and regress $q_{int} - f_{it}$ on $(r_{int})_{n=1\dots N}$, r^θ_{it}, r^ψ_{it} and obtain ξ_i and χ_i by OLS:

$$q_{int} - f_{it} = \sum_m \mu_{inm} r_{imt} + \xi_i r^\theta_{it} + \chi_i r^\psi_{it} + \nu^\perp_{int} \tag{21}$$

The identification comes from the variations of $\theta_{im,t-1}$ in r^ψ_{it} and r^θ_{it}.

In the final equation, we need the different asset classes to estimate μ_{inb}, yet ξ_i and χ_i are common across asset classes. We explore an alternative set of assumptions in Section 5.4 that allow us to estimate (ξ_i, χ_i) also using cross-sectional information.

5.4 Extension

In this section, we discuss an alternative set of assumptions on demand shocks that allow us to estimate ξ_i and χ_i using cross-sectional return data as well. In particular, we assume that demand shocks follow a factor model,

$$\nu_{int} = \sum_k \lambda_{ik} \eta_{nk} F_{kt} + u_{int},$$

where $\mathbb{E}[u_{int} u_{jms}] = 0$, except when $i = j$, $n = m$, and $s = t$. F_{kt} corresponds to the realization of the k-th factor, λ_{ik} is the sensitivity of investor i to to factor k, and η_{nk} is the loading of asset class n to factor k. Similar to assumption (18), we assume

$$F_{kt} = \sum_m \gamma^0_{mk} r_{mt} + F^\perp_{kt}.$$

This replaces the intercept from $\sum_m \mu_{inm} r_{mt} = \sum_m \mu^0_{inm} r_{mt} - (\xi_i + \chi_i \psi_n) r_{nt}$ to

$$\sum_k \lambda_{ik} \eta_{nk} \sum_m \gamma^0_{mk} r_{mt} - (\xi_i + \chi_i \psi_n) r_{nt}.$$

To gain intuition, we can look at a single factor ($K = 1$):

$$\lambda_i \eta_n \sum_m \gamma^0_m r_{mt} - (\xi_i + \chi_i \psi_n) r_{nt} = \lambda_i \eta_n \delta_t - (\xi_i + \chi_i \psi_n) r_{nt}.$$

The variation in r_{nt} (or r_{int} in case of investor-specific returns) that cannot be explained by $\lambda_i \eta_n \delta_t$ can then be used to identify ξ_i and χ_i as well. This is more powerful as it uses cross-sectional information in addition to only time-series information. The model can be estimated using nonlinear least squares.

5.5 Reaching for yield and implications for the Bernanke and Kuttner (2005) effect

We now show how the framework developed so far can be extended easily to capture reaching for yield incentives in a parsimonious manner.

Role of the risk-free rate Note that if there's a shock r_f to the risk-free rate, then the risk premium in (10) changes as

$$\frac{\Delta \pi_n}{\pi_n} = -\psi_n \left(r_n + T_n r_f \right), \tag{22}$$

where

$$T_n = \frac{1}{\delta_n + \phi_r}, \tag{23}$$

and where δ_n is the dividend yield of asset n and ϕ_r captures the mean-reversion of the short rate in continuous time. T_n is the "normative sensitivity" or "effective duration" of the price of asset n: it is a form of duration, and is lower when the interest rate movement mean-reverts faster. Then we have[18]

$$q_n = -\sum_m \left(1_{n=m} - \theta_m \right) \chi r_m - \sum_m \left(1_{n=m} - \theta_m \right) \xi \psi_m \left(r_m + T_m r_f \right)$$

which gives an extra term $-\zeta_{nf} r_f$ in the specification of q_n with

$$\zeta_{nf} = \xi \left(\psi_n T_n - \sum_m \theta_m \psi_m T_m \right) \tag{24}$$

Modeling "reaching for yield" We now focus on the role of the risk-free rate, r_f. We take it to be real, postponing issues of nominal effects (e.g. nominal illusion, or effects of inflation) to a future iteration. The demand for asset n is

$$q_n = -\zeta_{nf} r_f - \sum_m \zeta_{nm} r_m, \tag{25}$$

where ζ_{nf} is the semi-elasticity of the reaction to the risk-free rate r_f. The budget constraint forces

$$\sum_n \theta_n \zeta_{nf} = 0. \tag{26}$$

[18] We abstract from subscripts i and t for simplicity and assume $\nu_{int} = f_{it} = 0$.

Normatively, we should have (24) with $\xi = 1$. We propose to parametrize

$$\zeta_{nf} = \xi_f \left(\psi_n T_n - \sum_m \theta_m \psi_m T_m \right) + c \text{ for } n \neq f \tag{27}$$

where $c > 0$ indicates that if the risk-free rate is lower, the agent take more risk across the board, proportionally to the pre-existing portfolio, and ξ_f might be 0. When n is an asset different from the risk-free asset, c indicates a propensity to "chase the yield"; when $n = f$, $\zeta_{ff} \leq 0$ for the risk-free rate. By accounting, ζ_{ff} is given by (26), i.e.

$$\zeta_{ff} = -\frac{1 - \theta_f}{\theta_f} c. \tag{28}$$

Simple calibration To get a sense of c, let us consider the benchmark of an investor with a target real rate of return. The expected return is

$$R = r_f + \theta_r \bar{\pi}$$

where $\bar{\pi} = \frac{\sum_{n \neq f} \theta_n \pi_n}{\sum_{n \neq f} \theta_n}$ is the average risk premium on risky assets, and $\theta_r = \sum_{n \neq f} \theta_n$ is the share of risky assets. So, if the take R is constant, and $\bar{\pi}$ is constant (or, if $\xi_f = 0$), when r_f changes, we should have

$$d\theta_r = -\frac{1}{\bar{\pi}} dr_f$$

so $d \ln \theta_r = -\frac{1}{\theta_r \bar{\pi}} dr_f$. Hence, a pure "chasing the yield" investor exhibits:

$$c = \frac{1}{\theta_r \bar{\pi}} \tag{29}$$

With $\theta_r = 0.95$ and $\bar{\pi} = 3\%/year$, we obtain $c \simeq 30$ *years*. In general, we might expect a large, positive c.

The Bernanke and Kuttner (2005) effect and "reaching for yield" We conclude this section by showing that we need a "reaching for yield" force to account for the Bernanke-Kuttner effect, i.e. the fact that some risky asset (for instance, stocks) prices overreact to a change in the real rate.

To make the point, we consider a permanent change in the real rate. Take a model with just stocks and the short term bonds, and a permanent shock to r. The demand is

$$0 = -(1 - \theta_n) r_n - \kappa \delta_n (r_n + r_f T_n) - c r_f \tag{30}$$

where $T_n = \frac{1}{\delta_n}$ is the stock's duration, cr_f is the "chasing the yield" flow, and κ the sensitivity of

log portfolio weights to shocks to expected returns. So

$$r_n = -M_n^{BK} r_f T_n, \qquad M_n^{BK} = \frac{c\delta_n + \kappa\delta_n}{1 - \theta_n + \kappa\delta_n} \tag{31}$$

So, if there is no chasing the yield ($c = 0$), we get an underreaction of stock prices to changes in the risk-free rate ($M^{BK} < 1$). We have an overreaction of the stock price (as Bernanke-Kuttner found) iff $M^{BK} > 1$, i.e. iff $c\delta_n > 1 - \theta_n$

As $c \gg \kappa$ and $1 - \theta_n \gg \kappa\delta_n$, the "Bernanke-Kuttner" multiplier is

$$M^{BK} \simeq \frac{c\delta_n}{1 - \theta_n}$$

so if it is about 2, then we must have $c = M^{BK} \frac{1-\theta}{\delta} \simeq 2\frac{0.15}{0.03} = 10$.

The same idea would hold for long term bonds: one needs a "chasing the yield" element to generate an overreaction of the yield curve to movements of the short term rates. That said, the evidence for overreaction is clear for risky long term bonds (e.g. emerging market bonds) and for stocks, yet less clear for Treasuries.

5.6 Empirical results

Soon to be finalized.

6 Conclusion

We use new monthly security-level data on portfolio holdings and flows of U.S. households to estimate asset demand across asset classes and individual assets. Our data features broad coverage across the wealth distribution – including ultra-high-net-worth (UHNW) households – and spans multiple asset classes, covering both public and private assets.

Our data have two important advantages. First, we have data on UHNW individuals, with well over a thousand households who own more than $100 million in assets. This group of households that may be relevant for asset prices is typically under-represented in other data sources. The broad coverage across the wealth distribution also allows us to extrapolate our estimates to construct demand curves for the "representative US household." Second, we have broad coverage across asset classes and at high frequencies. The assets classes covered in the data include public and private assets and are all disaggregated to security-level data. Such a broad perspective is not even available for most US institutions.

We provide four main contributions. First, we provide two methodological contributions to analyze flows and asset demand across broad asset classes. We first develop a descriptive factor

model for flows to measure (i) how investors allocate capital to cash and to all risky asset classes combined and (ii) how investors reallocate capital across risky asset classes. Next, we develop a new methodology to identify and estimate demand curves across asset classes. This approach relies on economically-motivated restrictions on the covariance matrix of flows and returns, and can be applied in many settings where broad coverage of holdings is unavailable.

Second, our empirical results show that the flow to risky assets (and particularly equities) is pro-cyclical for less wealthy households (assets below $3 million) and counter-cyclical for wealthy households (assets above $10 million and in particular above $100 million). Wealthy households therefore stabilize fluctuations in risky asset markets, while less wealthy households contribute to markets' excess volatility.

Third, the factor model identifies three key rebalancing factors that are economically meaningful. These factors present bets on the long-term equity premium, the credit premium, the municipal bond premium, and the premium on US equities versus global equities.

Fourth, we estimate a demand model for the cross-section of US equities and find that demand elasticities for individual stocks decline in wealth and – in line with earlier literature – are generally very low.

References

Ameriks, John and Stephen P. Zeldes, "How do household portfolio shares vary with age?," Technical Report, National Bureau of Economic Research 2004.

Balloch, Cynthia Mei and Julian Richers, "Asset Allocation and Returns in the Portfolios of the Wealthy," 2021.

Barber, Brad M and Terrance Odean, "Trading is hazardous to your wealth: The common stock investment performance of individual investors," *The journal of Finance*, 2000, *55* (2), 773–806.

Bauer, Michael D and Eric T Swanson, "A reassessment of monetary policy surprises and high-frequency identification," Technical Report, National Bureau of Economic Research 2022.

Bender, Svetlana, James J Choi, Danielle Dyson, and Adriana Z Robertson, "Millionaires speak: What drives their personal investment decisions?," *Journal of Financial Economics*, 2022, *146* (1), 305–330.

Bernanke, Ben S and Kenneth N Kuttner, "What explains the stock market's reaction to Federal Reserve policy?," *The Journal of finance*, 2005, *60* (3), 1221–1257.

Betermier, Sebastien, Laurent E Calvet, Samuli Knüpfer, and Jens Kvaerner, "What Do the Portfolios of Individual Investors Reveal About the Cross-Section of Equity Returns?," *Available at SSRN 3795690*, 2022.

Bretscher, Lorenzo, Lukas Schmid, Ishita Sen, and Varun Sharma, "Institutional Corporate Bond Pricing," Technical Report, Swiss Finance Institute - HEC Lausanne 2022.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini, "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes," *Journal of Political Economy*, 2007, *115* (5), 707–747.

Calvet, Laurent E, John Y Campbell, and Paolo Sodini, "Fight or flight? Portfolio rebalancing by individual investors," *The Quarterly journal of economics*, 2009, *124* (1), 301–348.

Calvet, Laurent E, John Y Campbell, Francisco Gomes, and Paolo Sodini, "The cross-section of household preferences," Technical Report, National Bureau of Economic Research 2021.

Campbell, John Y, Tarun Ramadorai, and Benjamin Ranish, "Do the rich get richer in the stock market? Evidence from India," *American Economic Review: Insights*, 2019, *1* (2), 225–40.

Cremers, KJ Martijn and Antti Petajisto, "How active is your fund manager? A new measure that predicts performance," *The review of financial studies*, 2009, *22* (9), 3329–3365.

Davis, Steven J. and John Haltiwanger, "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *The Quarterly Journal of Economics*, 1992, *107* (3), 819–863.

Egan, Mark L, Alexander MacKay, and Hanbin Yang, "What Drives Variation in Investor Portfolios? Evidence from Retirement Plans," Technical Report, National Bureau of Economic Research 2021.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri, "Heterogeneity and persistence in returns to wealth," *Econometrica*, 2020, *88* (1), 115–170.

Fama, Eugene F. and Kenneth R. French, "A Five-Factor Asset Pricing Model," *Journal of Financial Economics*, 2015, *116* (1), 1–22.

Gabaix, Xavier and Ralph SJ Koijen, "Granular instrumental variables," Working Paper 28204, National Bureau of Economic Research December 2020.

Gabaix, Xavier and Ralph SJ Koijen, "In search of the origins of financial fluctuations: The inelastic markets hypothesis," Technical Report, National Bureau of Economic Research 2021.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, "Five facts about beliefs and portfolios," *American Economic Review*, 2021, *111* (5), 1481–1522.

Gürkaynak, Refet S, Brian Sack, and Eric Swanson, "The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models," *American economic review*, 2005, *95* (1), 425–436.

Haddad, Valentin, Paul Huebner, and Erik Loualiche, "How Competitive is the Stock Market? Theory, Evidence from Portfolios, and Implications for the Rise of Passive Investing," Technical Report, National Bureau of Economic Research 2022.

Hoopes, Jeffrey, Patrick Langetieg, Stefan Nagel, Daniel Reck, Joel Slemrod, and Bryan Stuart, "Who sold during the crash of 2008-9? evidence from tax-return data on daily sales of stock," Technical Report, National Bureau of Economic Research 2016.

Koijen, Ralph SJ and Motohiro Yogo, "A demand system approach to asset pricing," *Journal of Political Economy*, 2019, *127* (4), 1475–1515.

Koijen, Ralph SJ and Motohiro Yogo, "Exchange Rates and Asset Prices in a Global Demand System," 2020. NBER Working Paper 27342.

Koijen, Ralph SJ, Robert J Richmond, and Motohiro Yogo, "Which investors matter for equity valuations and expected returns?," *Working Paper*, 2021.

Nakamura, Emi and Jón Steinsson, "High-frequency identification of monetary non-neutrality: the information effect," *The Quarterly Journal of Economics*, 2018, *133* (3), 1283–1330.

Parker, Jonathan A., Antoinette Schoar, Allison Cole, and Duncan Simester, "Household Portfolios and Retirement Saving over the Life Cycle," Technical Report, National Bureau of Economic Research 2022.

Rigobon, Roberto, "Identification through Heteroskedasticity," *The Review of Economics and Statistics*, 2003, *85* (4), 777–792.

Appendix

A Literature review

In Table A1, we summarize related literature on portfolio choice decisions by households.

B Construction of Monetary Policy Shocks

B.1 Data construction

We construct monetary policy shocks using the first four Eurodollar future contracts with maturities in March, June, September and December traded on the Chicago Mercantile Exchange. For each contract, we compute the price change in a thirty minute window that starts ten minutes before and ends twenty minutes after each FOMC meeting. Specifically, we construct the price change by using the last trade that occurred more than ten minutes before teach FOMC meeting and the first trade that occurred more than twenty minutes after each FOMC meeting. In the peak of the COVID-19 crisis, the Fed held an emergency FOMC meeting at 5 p.m. EST on March, 15^{th} 2020 which was a Sunday. Because weekly trading on the Chicago Mercantile Exchange takes place from Sunday 5 p.m. CST to Friday 5.45 p.m. CST then, for this meeting only, we construct the price change for each future contract as the difference between the first price quoted on Sunday evening and the last price quoted on Friday evening. For the period before 2019, our list of FOMC meetings follows Bauer and Swanson (2022). For the period after 2019, we include in our list of FOMC meetings the release of four notation votes by the Fed which occurred on March 19^{th}, March 23^{rd}, March 31^{st} and August 27^{th} 2020. In constructing price changes, we restrict our dataset to trades with strictly positive volume as well as to trades that are flagged by the Chicago Mercantile Exchange and Tick Data as "normal trades". The second cleaning excludes trades such as block trades, privately negotiated trades, subtrades etc. which are more likely to depart from prevailing market prices. Monetary policy shocks are then constructed as the first principal component of the changes in Eurodollar futures prices around FOMC announcements. We rescale our measure of monetary policy shocks so that a one-unit change in the principal component corresponds to a one percentage point change in the price of the fourth Eurodollar future.

B.2 Replication of Bauer and Swanson (2022)

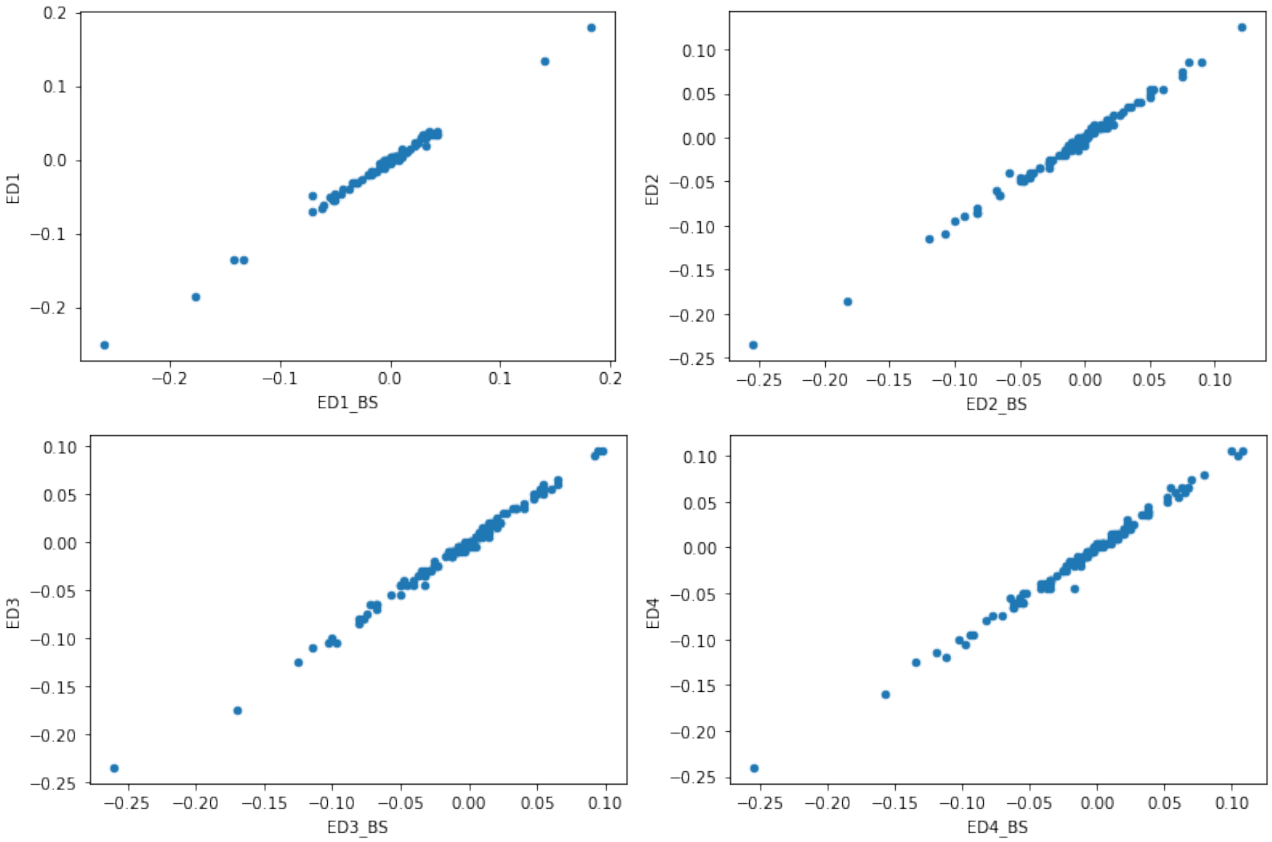
To assess the quality of the procedure that we use to construct MPS, we focus on the period 2005-2019 and scatter our reconstructed change in the prices of Eurodollar futures contracts at each

Table A1: **Summary of Literature on Household Portfolio Choice**

Source	Data Source	Coverage	Asset Classes	Key Questions
Barber and Odean (2000)	Brokerage firm	US (1991-1996)	Equity	Trading frequency and portfolio tilts of households
Giglio, Maggiori, Stroebel, and Utkus (2021)	Survey to Vanguard Clients	US (2017-2020)	Equity	Relationship between investor beliefs and portfolios
Bender, Choi, Dyson, and Robertson (2022)	Survey through UBS	US (March 2018)	Various	Drivers of individual investment decisions
Meeuwis, Parker, Schoar, and Simester (2018)	Financial Institution	US (2015-2017)	Equity, Fixed income, Alternatives	Relationship between investor beliefs and portfolios
Hoopes et al. (2017)	IRS	US (2008-2009)	Equity	Trading behavior during market distress
Balloch and Richers (2021)	Addepar	US (2016-2020)	Various	Return heterogeneity by wealth
Egan, MacKay, and Yang (2021)	BrightScope Beacon	US (2009-2019)	401(k) Holdings (Various)	Determinants of 401(k) allocations, focusing on risk aversion and beliefs
Campbell, Ramadorai, and Ranish (2019)	National Securities Depository Limited (NSDL)	India (2002-2011)	Equity	Relationship between return heterogeneity and equity wealth inequality
Calvet, Campbell and Sodini (2007)	Swedish Wealth and Income Registry	Sweden (1999-2002)	Various	Efficiency of household investment decisions focusing on under-diversification and non-participation
Calvet, Campbell and Sodini (2009)	Swedish Wealth and Income Registry	Sweden (1999-2002)	Various	Determinants of portfolio rebalancing and participation in risky financial markets
Calvet, Campbell, Gomes and Sodini (2021)	Swedish Wealth and Income Registry	Sweden (1999-2007)	Various	Distribution of preference parameters across households
Massa and Simonov (2006)	Longitudinal Individual Data for Sweden (LINDA)	Sweden (1995-2000)	Various	Portfolio allocation to hedge non-financial income
Fagereng, Guiso, Malacrino, and Pistaferri (2020)	Norwegian tax records	Norway (2004-2015)	Various	Return heterogeneity by wealth
Betermier, Calvet, Knupfer, and Kvaerner (2022)	Norwegian administrative data	Norway (1996-2017)	Equity	Relation between individual portfolios and cross-sectional equity returns
Dorn and Weber (2013)	German retail bank	Germany (2007-2011)	Equity	Portfolio allocation during market distress

This table summarizes the literature on household portfolio choice that is relevant for our work. For each source, we report the data source, the coverage (sample, location, and timeline), the main asset classes of interest, and key research questions addressed in the work.

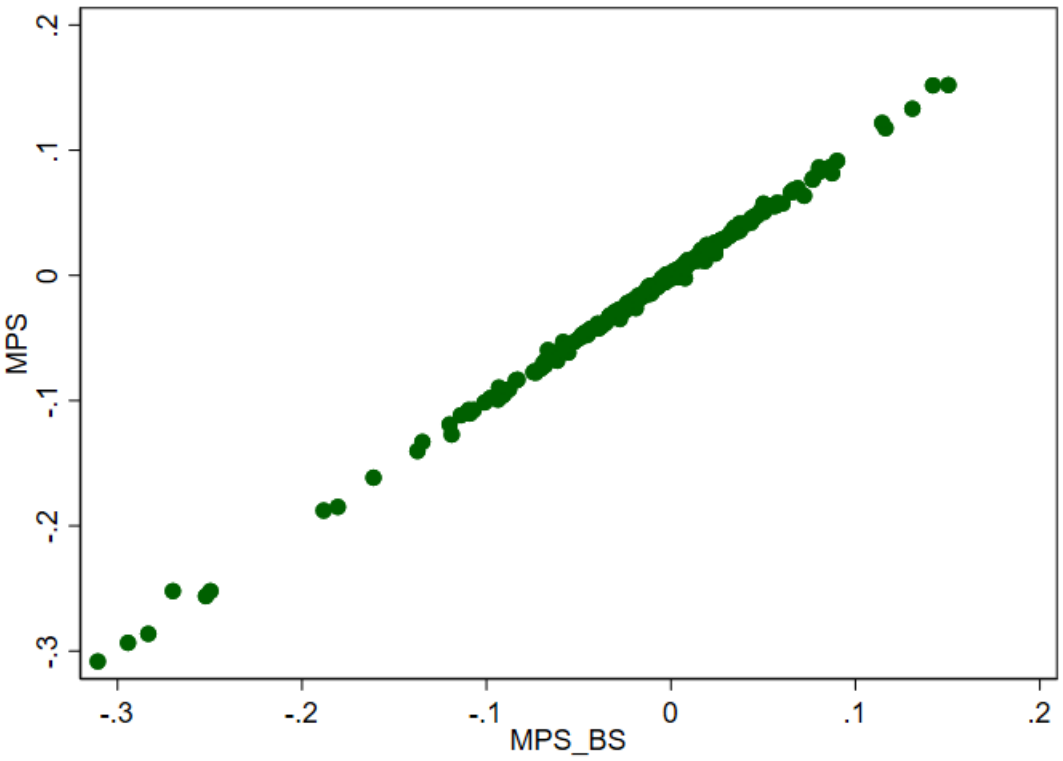
Figure A1: Change in Eurodollar futures contracts around FOMC meetings



FOMC meeting against the original measures in **?**. The correlations are 99.67% for ED1, 99.65% for ED2, 99.64% for ED3 and 99.61% for ED4.

As a second assessment, we consider our reconstructed changes in Eurodollar futures prices over the period 2005-2019 and reconstruct the MPS measure in Bauer and Swanson (2022) using data from 1988 to 2019. Below we scatter the MPS reconstructed using our data against their original MPS measure. The correlation is 99.9%.

Figure A2: Reconstructed MPS vs original MPS in Bauer and Swanson (2022)



Online Appendix

Xavier Gabaix Ralph S.J. Koijen Federico Mainardi

Sangmin S. Oh Motohiro Yogo

A Additional Details on Addepar Data

A.1 Data Structure

We have monthly data at security level on positions held and returns gained by individual investor accounts.

A.1.1 Variables

There are five classes of variables: (i) portfolio and security identifiers, (ii) firm identifiers, (iii) asset class and investment identifiers, (iv) holdings, flows and returns, and (v) variables related to other data sources. We next describe each in detail.

Portfolio and Security Identifiers We observe a unique identifier *portfolio_entity_id* for each account held by investors in our dataset. For securities held by investors, we observe four main identifiers. The first identifier *position_entity_id* is internally generated by Addepar and uniquely identifies a security. While *position_entity_id* is available for any security in the dataset, it is also complemented by CUSIP, ISIN and Sedol for securities for which these additional identifiers are available.

Firm Identifiers While we do not observe a unique identifier for firms/advisors, we observe a detail classification of firms based on the nature of their activities. From *firm_vertical*, any firm is first classified as Advisor, Broker Dealer, Consolidators, Family Office, Institutional, Other. Each broad classification in *firm_vertical* is further broken down into *firm_sub_vertical,* the details of which are summarized in Table A2.

Asset Class and Investment Identifiers The dataset spans a variety of asset classes. For each security, we observe the asset class entered by custodians/advisors in *input_asset_class*. Depending on the position, this input can be entered either manually or chosen from a precompiled list. *input_asset_class* is then processed internally by Addepar and rearranged into two additional classifications. The first one is *output_asset_class* which classifies any security in a broad asset class (e.g. Equities, Fixed Income). The second one is *sub_asset_class* that, for each broad asset class (e.g.

Table A2: **Firm Classification**

Category	Type
Advisor	Hybrid Registered Investment Advisor (Hybrid RIA), Independent Registered Investment Advisor (Independent RIA)
Broker Dealer	B/D Advisor, Bank Trust, National and Regional B/D, Private Bank, Wirehouse
Consolidators	Strategic Acquirer, Other
Family Office	Multi-Family Office, Single Family Office
Institutional	Foundation, Hedge Fund, Outsourced Chief Investment Officer (OCIO), Pension
Other	Fund Administrator, Software/Service Provider

Equities), classifies any security within a narrower asset class (e.g. US Equities, Global Equities). Separately from asset classes, we observe the type of investment associated to each position held by each investor. A broad classification is reported in investment_type. Within each broad classification in *investment_type*, we observe a narrower classification in *investment_sub_type*. Importantly, neither investment_type nor *investment_sub_type* are subsets of *sub_asset_class*. Indeed, two positions may have different sub_asset_class but same *investment_sub_type*.

Holdings, Flows, and Returns We also observe monthly holdings, flows, and returns for each position held by each investor. For each position, we observe dollar holdings at the beginning of the month in *starting_value* while dollar holdings at month-end are reported in *ending_value*. We observe a synthetic measure of monthly dollar flows in *net_cashflow* as well as the break down of *net_cashflow* into *buys* and *sells*. For specific asset classes, we separately observe measures of investment commitments made by the investors, contributions and distributions (*total_commitments_since_inception, total_commitments, total_contributions, unfunded_commitments, fund_distributions_and_dividends).* Turning to return measures, for each position held by each investor we observe monthly time-weighted return *twr*, internal rate of return *irr*, and dollar return *total_return*. We further observe the breakdown of gains into realized and unrealized, where unrealized gains refer to unsold positions.

Variables Related to Other Sources The dataset further includes variables from alternative data sources. From Preqin, we observe *preqin_id, vintage*, *strategy* and *substrategy*. All variables are also included in the Preqin manual where *preqin_id* is called *FUND ID*, *vintage* is called *VINTAGE / INCEPTION YEAR*, *strategy* is called *ASSET CLASS* and *sub_startegy* is called *STRATEGY*. Using *preqin_id* we can then merge all information in the Preqin manual into the main dataset. From Morningstar, we observe *morningstar_asset_class*, *morningstar_us_asset_class*, *morningstar_global_asset_class*, *morningstar_business_country_cla*, *morningstar_region_breakdown*,

morningstar_category, *morningstar_security_type*, *morningstar_industry*. From SIX, we observe *six_instrument_type*, *six_security_type*, *six_domicile2*. From Pitchbook and HFRI, we observe *pitchbook_id* and *hfri_id* respectively. We observe a separate classification for bonds in *sp_bond_type*, *sp_bond_sub_type* and *sp_bond_domicile_of_issuer*. Finally, we observe three additional identifiers internally produced by Addepar, namely *issuer_id*, *security_id* and *model_type*. The latter is mainly used as an input in Addepar Navigator to produce predictions about prices and volumes.

A.2 Asset Class and Investment Type Taxonomy

Each position in the data is associated with an asset class and an investment type. The asset class represents a classification of the position into a more general asset category. The investment type is independent of the asset class and refers to the nature of positions held by investors. For instance, a position in a common stock would have asset class equal to Equities and investment type equal to Common Equity. A position in an equity mutual fund would have asset class equal to Equities but investment type equal to Mutual Funds. In Table A3 and A4, we provide the breakdown of asset classes into sub asset classes as well as of investment types into sub investment types.

1. Cash and Cash Equivalents "Cash" includes cash held for non-investment/hedging purposes for personal liquidity/working capital or savings. Examples include Checking accounts, savings accounts, and FX trading accounts. "Cash Equivalents" includes Commercial Paper, Repos, T-Bills, Variable Rate Demand Notes (VRDNs) and obligations (VRDOs), Auction Rate Securities (ARS), and Money Market Funds (MMFs).

2. Fixed Income "US Municipals/Tax Exempt" includes US Municipal (General Obligation / Revenue) bonds such as Sewer Bonds, School Bonds, GO City/State/Town bonds. "US Treasuries and Agencies" refers to US Government and Agency (Fannie/Freddie/Ginnie) debt such as US 30-year or GNMA 10-year. "US Tips" refers to Treasury Inflation-Protected Securities (TIPS). "US Investment Grade Credit" refers to BBB And above-rated debt and "US High Yield" refers to Below BBB-rated debt. "US Bank Loans" refers to traded bank loans such as term loan bullet and balloons. "International Developed Markets" refers to corporate and government debt securities and pooled vehicles with primary exposure to developed markets, while "Emerging Markets" refers to similar assets with primary exposure to emerging markets. "Opportunistic" refers to unconstrained / total return bond strategies. "Unknown Fixed Income" contains assets that client tagged as "Fixed Income" without more granular information, and "Other Fixed Income" contains all the remaining categories.

3. Equities "US Equities refers to equities issued in the US, and "Developed Markets - Americas" refers to those issued by primarily US-domiciled or Canadian companies. "Developed Markets

- EMEA" refers to equities from France, Germany, Switzerland, South Africa, etc., while "Developed Markets - Asia Pacific" refers to equities from Australia, Japan, Hong Kong, Singapore, New Zealand. "Emerging & Frontier Markets" refers to equities from China, Brazil, Russia, Venezuela, Romania, Vietnam, etc. "Global Equities" refers to funds that span multiple markets. "Concentrated Equity Positions" contain positions with >$1B in US companies. "Unknown equities" include assets that are identified as equities, but with no other information available, and "Other Equities" contains all the remaining categories.

4. Alternatives "Hedge Funds" includes pooled NAV vehicle as well as asset management products but not regulated like a mutual fund. The funds encompass a range of strategies and can invest in a wide range of securities and asset types. "Private Equity & Venture" refers to pooled drawdown vehicles investing in a range of underlying assets, including buyout, growth equity companies, credit, infrastructure funds, and venture funds. Only includes pooled vehicles, not direct investments. "Real Estate Funds" refers to pooled vehicles investing in real estate excluding REITs. "Direct Private Companies" refers to equity in private companies, including VC-backed companies or "regular operating companies" like hardware stores, gas stations, etc. "Direct Real Estate" refers to ownership or ownership stake in commercial or residential real estate. "Direct Loans" refers to non-securities debt investments, direct loans (lending). "Concentrated Alts Positions" refers to single positions with > $1B in value. "Unknown Alts" refers to alternatives funds of unknown classes, while "Other Alts" includes physical or derivative exposure to commodities, Traded / untraded crypto positions, and other pooled vehicles including exotic pooled vehicles.

5. Non-Financial Assets "Collectibles and Other" includes all remaining assets that do not fit into any of the narrow asset classes.

6. Liability "Liability" refers to Loans taken out by the client (as opposed to loans made by the client, which are considered assets).

7. Mixed Allocation "Asset Allocation Vehicle" refers to Multi asset class vehicles. "Held Away Accounts" refers to Single line positions where we don't know what the underlying investments are, generally only the type of account (eg 401k, which is an account structure, but can hold a lot of different asset types).

Table A3: **Asset Class Taxonomy**

Broad Asset Class	Narrow Asset Class
Alternatives	Direct Loans, Direct Private Companies, Direct Real Estate, Hedge Funds, Other Alts, Private Equity & Venture, Real Estate Funds, Unknown
Cash and Cash Equivalents	Cash, Cash Equivalents
Equities	Developed Markets - Americas, Developed Markets - Asia Pacific, Developed Markets - EMEA, Emerging & Frontier Markets, Global Equities, US Equities, Other, Unknown
Fixed Income	Emerging Markets, International Developed Markets, Opportunistic, US Bank Loans, US High Yield, US Investment Grade Credit, US Municipals/Tax Exempt, US TIPS, US Treasuries and Agencies, Other, Unknown
Mixed Allocation	Asset Allocation Vehicle, Held Away Accounts
Non-Financial Assets	Collectibles and Other

Table A4: **Investment Type Taxonomy**

Category	Type
Bank/Brokerage Account	Brokerage/FX Cash Account, US Bank Account
Collectibles	Collectibles
Derivative	Listed Option
Equity	American Depository Receipts (ADR), Common Equity, International, Preferred Equity, Restricted Equity, Rights/Warrants, Other
Fixed Income	ABS/MBS, Certificate of Deposit (CD), Corporate Bonds, Muni Bonds, Treasuries, US Agency, Other
Held Away	Employee Benefit Plan, Managed Account, Tax-Advantaged Plan, Other
Insurance	Annuities, Other
Limited Partnership	Drawdown LP, NAV LP, Unknown
Loans	Mortgage, Security-Based Loan (SBL) / Margin Loan, Unsecured, Other
Other	Crypto, Other
Private Company	Operating Company, Private Option, Venture Backed Company
Public Fund	Closed End Fund, ETF, Master Limited Partnership (MLP), Money Market Fund (MMF), Mutual Fund, REIT, Other
Real Estate	Commercial Real Estate, Residential Real Estate, Unknown