Do Household Biases Matter for Asset Prices?

Carter Davis, Samuli Knüpfer, Jens Soerlie Kvaerner, Bahar Sen-Dogan, Petra Vokata*

November 12, 2026

Abstract

We apply a demand system to complete ownership records for Norwegian stocks to quantify the importance of household biases for asset prices. Households own less than one-fifth of the market but account for up to one-third of cross-sectional return variance. Biased households—those scoring high on empirical measures of behavioral biases—account for most of this variance. Debiasing households by relocating 20% of their wealth from the most- to the least-biased households compresses anomaly spreads by 130 basis points a year, or 27% of the anomaly spread, on average. Household biases therefore matter for both price variation and expected returns in equilibrium.

*Davis: Fisher College of Business, The Ohio State University, 2100 Neil Avenue, Room 828, Columbus, OH 43210; carter.davis@fisher.osu.edu. Knüpfer: Aalto University School of Business, Ekonominaukio 1, 02150 Espoo, Finland, and BI Norwegian Business School and Research Institute of Industrial Economics (IFN); samuli.knupfer@aalto.fi. Kvaerner: Tilburg University, Warandelaan 2 5037 AB Tilburg, Netherlands; jkverner@gmail.com. Sen-Dogan: Tilburg University, Warandelaan 2 5037 AB Tilburg, Netherlands; b.sendogan@tilburguniversity.edu. Vokata: Fisher College of Business, The Ohio State University, 2100 Neil Avenue, Room 836, Columbus, OH 43210, and CEPR; vokata.1@osu.edu. We thank seminar and conference participants at the EFA, BI Norwegian Business School, CUNEF University, NFA, CEPR Household Finance Conference, Leibniz University Hannover, MFA, Norwegian School of Economics, Oslo Metropolitan University, The Ohio State University, Tilburg University, and Nick Barberis, Justin Birru, Per Bye, Aditya Chaudhry, Andrei Goncalves, Michael Haliassos, Paul Huebner, Ralph Koijen, Fulin Li, Terry Odean, and René Stulz for discussions and comments. We gratefully acknowledge financial support from Finansmarkedsfondet at Research Council of Norway, Project No. 309855.

1 Introduction

How much do behavioral biases matter for asset prices? Since De Bondt and Thaler (1985), behavioral biases have been offered as explanations for anomalies in stock returns. That biased investors can move prices when arbitrage is limited is now broadly accepted (see Barberis, 2018, for a survey). The magnitude of that price impact, however, remains an open question. A skeptic might say that irrational investors, typically households, are small, so their imprint on prices must be small as well.

We quantify the importance of household biases by applying a demand-system analysis in the spirit of Koijen and Yogo (2019) to granular data on all Norwegian investors.[1] We observe the complete ownership structure of listed equities in Norway, including households, institutions, and other investor sectors, and we use these data to unpack the behavioral origins behind investor demand. The combination allows us to quantify an equilibrium link between empirical measures of household biases and stock returns. The counterfactual exercises reveal that biased households, those scoring high on empirical measures of behavioral biases, account for a disproportionate fraction of cross-sectional variance. Debiasing households by relocating 20% of household wealth from the most- to the least-biased households compresses long–short anomaly spreads by 130 basis points a year, or 27% of the average anomaly spread.[2]

Our data allow us to construct empirical measures of household behavioral biases, an exercise we would not be able to perform with typical data on U.S. institutions. The premise underlying our analysis, which is supported by a vast household-finance literature[3], is that household portfolio choice and trading behavior are shaped by behavioral biases. We cannot observe those biases directly. Instead, we construct empirical measures of behavioral biases

[1] Related work shows that investor demand and trading structure move prices in settings including retail trading (Barber, Odean, and Zhu, 2008; Van der Beck, Jaunin, and Cohen, 2026; Li, 2025), index-level trading (Lochstoer and Muir, 2026), and price-impact bounds and disagreement (van der Beck, Bretscher, and Fu, 2025). We add an equilibrium counterfactual for measured household portfolio tilts, allowing every observed investor sector to respond.

[2] The counterfactual measures the price impact transmitted through household portfolios. Institutional sectors enter the equilibrium absorption system, but we do not classify institutions by the household bias measures. If institutions exhibit the same mechanisms, their economywide effects may be larger than the household channel we estimate.

[3] For surveys, see Campbell (2006), Barber and Odean (2013), Barberis (2018), Gomes, Haliassos, and Ramadorai (2021), and Campbell and Ramadorai (2025).

that follow that literature. We start with the three leading frameworks in Barberis (2018) (overconfidence, prospect theory, and extrapolation) and map them to four empirical measures observable in holdings and trades. Overconfidence maps to portfolio turnover; prospect theory maps to a tilt toward lottery-type stocks and to the disposition effect; and extrapolation maps to a measure of extrapolation based on past returns. Throughout the paper, we use "biased households" and "household biases" to refer to these empirical classifications. We sort households into bins by each score within wealth brackets, separating the measured bias from wealth.

Our administrative data cover direct holdings over 160 months, from January 2007 through April 2020. For households, we observe both direct holdings and ownership through holding companies such as family offices and family firms, especially common for more affluent investors. Households as a whole hold 18% of listed equity on the Oslo Stock Exchange, compared with 39% for institutions, 22% for the government, and 21% for other investors.

To turn household biases into asset-pricing magnitudes, we need a framework that maps holdings into equilibrium prices and expected returns. We use demand-system asset pricing (DSAP), which estimates investors' demand curves from holdings data (Koijen and Yogo, 2019; Koijen, Richmond, and Yogo, 2022). Each portfolio reveals desired holdings at prevailing prices. With instrumented price variation, the system estimates how those holdings respond to price changes. Aggregating demand curves and imposing market clearing allows us to invert the system: a counterfactual flow is translated into the price changes required for all investors to absorb it. This inversion lets us reallocate wealth from the most- to the least-biased households, let all investor sectors respond through their estimated demand curves, and map the resulting price distortions into next-month expected returns through pass-through estimates (Davis, Han, Sokolinski, and Tamoni, 2025a). Our implementation does not impose standard logit demand, in which one parameter governs all own- and cross-price substitution. Instead, we estimate a log-linear demand system that can be derived from CARA or CRRA portfolio choice, allows demand to depend on expected returns and covariances, and uses the factor-spillover structure of Haddad, He, Huebner, Kondor, and

Loualiche (2025) for substitution across stocks.[4]

In our first set of counterfactual exercises, we ask how much cross-sectional return variance is attributable to households and to their biases. Overall, households explain 33.4% of equal-weighted and 22.0% of value-weighted modeled variance despite holding only 18% of inside AUM. Relative to ownership, that is 1.9 times their AUM share under equal weighting and 1.2 times under value weighting. Institutions, by contrast, explain less variance than their AUM share.

Under hypothetical 50% AUM withdrawals from each classified household bin, with other demand shocks set to zero, the price-variance contribution is concentrated among the most-biased and the wealthiest households. The signed Shapley decomposition, which divides covariance equally between groups, assigns 45–85% of the equal-weighted variance to most-biased households in the top wealth decile, compared with only 2–8% for their counterparts in the remaining nine deciles combined. This concentration reflects two facts. First, the most-biased groups matter disproportionately in the scaled comparison that combines their bin variance-to-AUM ratio with the broad household sector's historical ratio: the resulting ratios range from 2.1–4.6 under equal weighting and 1.3–2.1 under value weighting. Second, household wealth itself is highly concentrated: the top decile holds almost 90% of household AUM, similar to the household wealth concentration in the U.S.

Taken together, these results imply that households, and biased households in particular, are far from price-irrelevant noise: their demand movements are large enough, and sufficiently distinct across stocks, to account for a meaningful component of equilibrium price variation.

In our second counterfactual exercise, we quantify the role of behavioral biases in driving anomaly returns, the central focus of many behavioral asset-pricing models. We start by collecting predictions from the literature on which anomalies may be driven by which biases. For example, Daniel, Hirshleifer, and Subrahmanyam (1998) argue that overconfidence can overprice glamour stocks, those with low fundamental-to-price ratios, and thereby generate value premia. Barberis, Jin, and Wang (2021) show that prospect-theory preferences can help explain a majority of 23 prominent anomalies through stocks' volatility, skewness, and

[4]Related work shows that these substitution patterns are central for demand multipliers when securities have heterogeneous substitutes (Chaudhary, Fu, and Li, 2023).

capital-gain overhang. And Grinblatt and Han (2005) argue that the disposition effect can explain momentum.

To test these predictions, we collect 139 anomaly portfolios from the global factor database of Jensen, Kelly, and Pedersen (2023) that are available in Norway. We map each prediction to the anomalies that fall under the proposed mechanism. For example, we use the ten fundamental-to-price anomalies to test the overconfidence–value prediction of Daniel et al. (1998). In total, this yields 18 bias–category predictions spanning 51 distinct anomalies.

In the counterfactual experiment, we assess the role of biases in driving anomaly returns by relocating 20% of household AUM from the most- to the least-biased households within wealth deciles, holding nonhousehold demand curves fixed and tracing the resulting price and expected-return responses through the demand system.

We find that 16 of 18 predictions have the expected positive sign, and 10 remain significant at the 5% level after accounting for multiple testing. On average, the most-biased households contribute 130 basis points a year to these anomaly spreads, or 27% of the mean realized spread. Our counterfactual exercises therefore imply that household biases matter quantitatively both for cross-sectional variance and for expected anomaly returns in equilibrium.

We build the counterfactual around recent DSAP innovations that address the main critiques of using estimated demand curves for policy-style counterfactuals. The common concern is Lucas-critique-style: the demand response estimated from observed price variation may not be the same response that applies to our household-flow shock. We address three versions. First, counterfactual flows may generate different cross-stock spillovers than the identifying price movements, a problem that is especially sharp in logit demand (Fuchs, Fukuda, and Neuhann, 2023); our baseline uses flexible factor spillovers of Haddad et al. (2025), and we also consider alternative spillover systems. Second, the counterfactual price response may have different persistence dynamics, which changes pass-through from price variation to expected returns (Davis, Kargar, and Li, 2025b; Davis, Kargar, Li, and Silva, 2026; van Binsbergen, David, and Opp, 2025); we estimate price persistence and expected-return pass-through separately, and the results are robust across a reasonable range of dynamic pass-through responses. Third, the shock may change perceived risk, so the relevant

counterfactual elasticity may differ from the baseline total elasticity (He, Kondor, and Li, 2025; Davis et al., 2026); we recover the expected-return and risk components of total elasticity and recompute the headline exercises under alternative endogenous-risk responses. Thus the baseline counterfactual holds preferences, characteristics, and estimated local elasticities fixed, while the robustness exercises ask whether the conclusions survive the main ways in which the relevant counterfactual demand curve might differ from the one estimated in the data.

Literature review. We contribute to three strands of the literature. The first is the literature on behavioral biases and the cross-section of stock returns. We provide a quantitative equilibrium link between household-level behavioral biases and anomaly returns by combining direct measurement of behavioral tilts with a structural counterfactual. Our results complement quantitative behavioral asset-pricing models (e.g., Barberis, Shleifer, and Vishny, 1998a; Barberis, Shleifer, and Wurgler, 2005; Barberis, Mukherjee, and Wang, 2016; Barberis et al., 2021; Barberis, 2018), which predict the characteristics that earn anomalous returns; we add evidence from investor portfolios, quantifying the equilibrium effect of household behavioral biases on expected anomaly returns. Our results complement reduced-form evidence on correlations between retail order flow and returns (e.g., Kaniel, Saar, and Titman, 2008; Kelley and Tetlock, 2013; Barber et al., 2008; Boehmer, Jones, Zhang, and Zhang, 2021; Frazzini and Lamont, 2008), between bias-related stock characteristics and returns (e.g., Kumar, 2009; Bali, Cakici, and Whitelaw, 2011; Frazzini, 2006), and between investor sentiment and anomalies (e.g., Baker and Wurgler, 2006; Stambaugh, Yu, and Yuan, 2012, 2015), by delivering the equilibrium magnitude from a structural counterfactual.

Financial economists have long debated whether stock-market anomalies reflect rational risk compensation or behavioral mispricing (Fama, 1965; Barberis and Thaler, 2003). The behavioral view holds that systematic investor biases can survive aggregation and affect equilibrium prices (Shleifer, 2000). Correlational evidence alone does not identify the equilibrium price impact of biased demand, since rational investors may absorb biased positions at unchanged prices. Moreover, Kozak, Nagel, and Santosh (2018) show that any returns-based test, whether of factor models or of characteristics, is observationally equivalent between rational and behavioral explanations, and argue that distinguishing the two requires structural

models with explicit assumptions about investor beliefs and preferences. We contribute a complementary structural counterfactual: using complete ownership data, we show that the return contributions of measured behavioral biases align with theoretical predictions across a broad anomaly universe.

The second contribution is to the literature on household finance. We are the first to unpack the household sector inside a demand-system framework, going beyond the residual-from-13F treatment of prior demand-system work. Nordic and U.S. household-finance research (e.g., Barber and Odean, 2001; Calvet, Campbell, and Sodini, 2007; Grinblatt, Keloharju, and Linnainmaa, 2012; Balasubramaniam, Campbell, Ramadorai, and Ranish, 2023; Betermier, Calvet, Knüpfer, and Kvaerner, 2024) has documented heterogeneity in household portfolios, returns, and behavioral profiles. We move from welfare and portfolio-choice questions to the equilibrium price-formation question, asking whether and how much that heterogeneity matters for the cross-section of stock returns. We show that household finance matters not only for welfare (Bhamra and Uppal, 2019), but also for asset prices.

The third contribution is to the literature on demand-based asset pricing. We extend the demand-system asset-pricing framework (e.g., Koijen and Yogo, 2019; Koijen et al., 2022; Gabaix, Koijen, Mainardi, Oh, and Yogo, 2022; Li, Sokolinski, and Tamoni, 2026; Gabaix, Gopikrishnan, Plerou, and Stanley, 2006; Campbell, Ramadorai, and Schwartz, 2009) along two dimensions. We add investor-level granularity and behavioral-bias sorts beyond the 13F universe. And we not only apply Haddad et al. (2025) but also show how to run counterfactuals in a way that is robust to concerns about cross-asset spillovers, dynamics, and endogenous risk. This latter contribution is most similar to Davis et al. (2026), but we are the first to do it at the investor level, in a simple and tractable way that addresses recent concerns about demand-system asset pricing. The richer system preserves the central object, the inverse-of-total-elasticity multiplier, while admitting more realistic substitution across stocks and allowing robustness to variation in shock persistence and risk effects.

The remainder proceeds as follows. Section 2 describes the data, the bias measures, and the anomaly portfolios. Section 3 develops the demand system and the counterfactual experiments. Section 4 reports the main results from the two counterfactual exercises, along with the robustness exercises. Section 5 concludes.

2 Data

We merge several administrative datasets to provide complete ownership records for all publicly listed Norwegian stocks. In what follows, we describe the data sources, the classification of investors and assets, the construction of household-level behavioral bias measures, and the anomaly portfolios used in counterfactual experiments.

2.1 Ownership Data

Norwegian Central Securities Depository. The Norwegian Central Securities Depository (Verdipapirsentral, hereafter VPS) records ownership of every share on the Oslo Stock Exchange (OSE). OSE is Norway's only regulated securities market, medium-sized by European standards (Doeskeland and Hvide, 2011). VPS is its sole depository. Each month, we observe every trading account; holdings sum to 100% of shares outstanding. By contrast, the standard approach in the demand-system literature uses 13F institutional filings and models households as the residual sector, which precludes modeling household demand on a more granular level. Koijen and Yogo (2019) therefore advocate Nordic administrative data; VPS is well suited to that agenda because ownership is complete at monthly frequency, with no residual sector. That completeness is also central to the counterfactual design: all investors affected by a wealth redistribution are observed.

Investor types. We assign each VPS trading account to one of seven investor types in two steps. First, we read VPS's own sector classification at the account's first appearance in the data and hold that assignment fixed thereafter.[5] Second, within each investor type we apply finer rules described below and in Section 2.2 for households. The seven types are Households, Institutions (mutual funds, pension funds, insurance companies), Governments, Banks, Foreign investors, Listed firms, and Non-listed firms.[6]

Households hold stocks directly and through private companies such as family offices. To capture both channels, we link VPS records to the Norwegian Tax Administration's

[5]The first-appearance rule ensures account stability over time, even if VPS occasionally reclassifies it.

[6]The Foreign category combines non-institutional foreign accounts with accounts of Norwegian residents that cannot be merged with Statistics Norway records. Household refers to an individual identified by a Social Security number.

Shareholder Register, which records ownership of every privately held Norwegian company. We trace household ownership through up to five layers of corporate ownership. For example, if 50% of company A belongs to an individual and the remaining 50% to company B, itself equally owned by two individuals, we assign company A to three households with shares of 50%, 25%, and 25%. We classify a privately held company's account as "household" if its ten largest household owners hold more than 50% of the company in any sample year.

Appendix J.2 validates this classification. Figure 8 shows that, within each wealth group, direct households and family offices hold the same average number of stocks per account. Table 24 shows that the cosine similarity between their average portfolio weights exceeds 0.80 in groups W3–W10, where both samples are large.

Stock characteristics. We collect stock prices, dividends, corporate events, shares outstanding, book equity, and net income from Titlon, a financial database for Norwegian universities. ESG scores come from MSCI. Four characteristics enter the demand system: CAPM beta (Beta), book equity, ESG greenness (ESG), and return on equity (ROE). Beta is the slope from daily regressions of stock returns on an equally weighted market index over a rolling 24-month window, with a Dimson (1979) adjustment using one lead and one lag of the market return to correct for asynchronous trading; we require at least 180 valid daily observations per window. When the Beta estimate is unavailable, we set Beta to one. We set negative book equity values to missing and carry forward the most recent observation. ROE equals net income divided by lagged book equity, both lagged by 12 months. We winsorize Beta and ROE at the 5th and 95th percentiles. We forward-fill missing ROE within stock and impute remaining gaps with the cross-sectional median that month. ESG greenness follows Pastor, Stambaugh, and Taylor (2022):

$$G_{n,t} = -(10 - Escore_{n,t}) \times \frac{Eweight_{n,t}}{100}, \tag{1}$$

where $Escore_{n,t}$ is the distance from a perfect environmental score and $Eweight_{n,t}$ is the firm's exposure to brown industries. The minus sign makes $G_{n,t}$ rise with greenness. When firm-level scores are missing, we use the IVA industry median. We standardize $G_{n,t}$ cross-sectionally to zero mean and unit variance each month.

Inside and outside assets. We classify each security as "inside" or "outside". Inside

assets are the stocks in the Betermier et al. (2024) sample, plus post-2018 listings that meet the same liquidity and coverage criteria. We drop stock-months in which the Titlon and VPS totals for shares outstanding differ by more than 5%. Outside assets are the remaining listed equities, together with bond and money market fund holdings.[7] The inside-asset universe averages about 120 stocks per month.

Figure 1 plots ownership shares over time by investor type. Institutions own 39% of the market, governments 22%, the household sector 18%, and foreign investors 16%; banks, listed firms, and non-listed firms hold the remaining 5%.

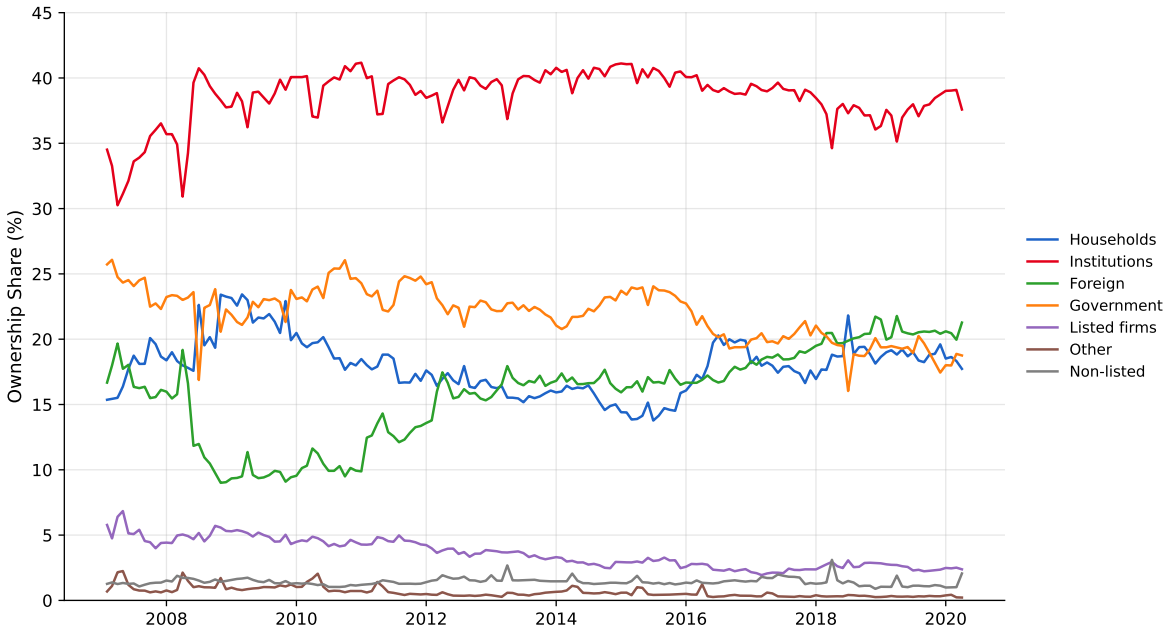


Figure 1. Ownership Shares of Norwegian Listed Equity by Investor Type.
This figure plots monthly ownership shares of the Norwegian inside-asset universe by investor type from January 2007 through April 2020. For each investor type and month, the ownership share equals the market value of all holdings in that investor type divided by the total market value of inside-asset stocks. Holdings are aggregated across all VPS accounts assigned to the investor type. Households include both direct household accounts and accounts attributed to households through family offices. Institutions include mutual funds, pension funds, and insurance companies. Governments are dominated by Folketrygdfondet, the domestic mandate of the Norwegian Government Pension Fund. Foreign investors are non-resident investors directly registered with VPS.

Table 1 summarizes the household-sector panel. The full sample contains 808,366 unique accounts over 160 months; 770,937 are direct household accounts and 37,429 are family-office accounts identified through the layered tax-registry trace. The median month observes

[7]Fund holdings come from Titlon, classified using the Thomson Reuters Lipper scheme. We exclude equity funds to avoid double-counting stocks already in the VPS ownership data.

394,098 active households, the peak month 439,346. Roughly 37% of households appear in at least 120 months. The long panel allows us to track each household's investment style over many years.

Table 1. Coverage of the Norwegian Household Holdings Panel

Sample period	
Panel period	Jan 2007 – Apr 2020
Number of months	160
Mean inside-asset stocks per month	120
Household coverage	
Unique households	808,366
Direct households	770,937
Family offices	37,429
Active households, median month	394,098
Active households, peak month	439,346
Share with ≥ 120 months in panel	37.1%

This table reports summary statistics for the household-sector holdings panel constructed from the Norwegian Central Securities Depository (VPS) and linked to the Norwegian Tax Administration's Shareholder Register. Direct households are individual VPS accounts. Family offices are corporate VPS accounts traced back to households. Active households are accounts that hold at least one inside-asset in a given month. Inside assets are stocks satisfying the liquidity and coverage criteria described in Section 2.1. The mean stock count is the time-series average across the sample.

2.2 Measuring Behavioral Biases

Complete ownership records, demographic identifiers, and monthly positions let us construct empirical measures of household behavioral biases. We cannot observe those biases directly. Following the household-finance and behavioral-asset-pricing literatures, we start from three leading behavioral frameworks: overconfidence, prospect theory, and extrapolation. As Barberis (2018) summarizes, these frameworks have been the main focus of the behavioral asset-pricing literature and help explain a wide range of facts about trading volume, asset-market fluctuations, and the cross-section of average returns. We then map the frameworks to four empirical measures that we can construct from holdings and trades. Overconfidence maps to portfolio turnover; prospect theory maps to a tilt toward lottery-type stocks and to the disposition effect; and extrapolation maps to a measure of extrapolation based on past returns. For each measure, we compute a household-level intensity score from lagged trading or holdings and sort households within wealth brackets for the counterfactual

experiments.

These empirical measures are not themselves the underlying biases. The disposition effect, for example, is an empirical trading pattern, a greater propensity to sell winners than losers, not a preference or belief; prospect theory is one of the leading explanations for that pattern (Shefrin and Statman, 1985; Barberis and Xiong, 2009; Barberis, 2018). In the same spirit, turnover, lottery tilts, and extrapolation are measured behaviors that the literature has linked to overconfidence, probability weighting and related lottery demand, and extrapolative beliefs, respectively. Throughout the paper, "biased households" and "household biases" refer to these empirical classifications.

General Framework

For each bias b, we compute a household-level intensity score $m^b_{i,t}$ from lagged trading behavior or portfolio holdings. Let $w_{i,n,t}$ denote household i's portfolio weight in stock n at month t, and let $A_{i,t}$ denote the household's total stock wealth in VPS. We sort households into 10 deciles of lagged total stock wealth ($A_{i,t-1}$) and, within each decile, into 5 bias-intensity quintiles, producing 50 household portfolios per bias. Quintile $q = 1$ contains the least biased households and $q = 5$ the most biased. The one exception is overconfidence: because portfolio turnover is heavily right-skewed in the household population and a substantial mass of households does not trade in a given month, we sort overconfidence into 3 bins (T1, T2, T3) within each wealth bracket and treat T3 as the most biased. Households with zero turnover in the lookback window are placed in T1. We compute each bias score only for households that hold or trade stocks during the period the score is based on.

1. Overconfidence

Overconfident investors trade too much. Barber and Odean (2001) show that excessive trading, driven by overconfidence, reduces net returns, and that the effect is strongest among men. In asset-pricing models, overconfidence and related belief-based mechanisms can affect prices through several channels: biased self-attribution can generate momentum and reversal; disagreement under short-sale constraints can generate speculative overpricing; sticky expectations can generate profitability predictability; and extrapolation can generate value-

related effects (Barberis et al., 1998a; Daniel et al., 1998; Hong and Stein, 1999; Hong and Sraer, 2016; Scheinkman and Xiong, 2003; Bouchaud, Krueger, Landier, and Thesmar, 2019; Miller, 1977; Diether, Malloy, and Scherbina, 2002; Lakonishok, Shleifer, and Vishny, 1994). We measure overconfidence with portfolio turnover (Odean, 1999; Barber and Odean, 2001), defined as the absolute deviation of the realized post-rebalancing weight from the passive weight that would obtain absent any trading:

$$\text{Turnover}_{i,t} = \frac{1}{2}\sum_{n}\left|w_{i,n,t} - w_{i,n,t}^{\text{passive}}\right|, \qquad w_{i,n,t}^{\text{passive}} = \frac{w_{i,n,t-1}\,(1+R_{n,t})}{1+R_{i,t}^{P}}, \tag{2}$$

where $R_{n,t}$ is stock n's month-t return and $R_{i,t}^{P}$ is household i's portfolio return. Subtracting $w_{i,n,t}^{\text{passive}}$ rather than $w_{i,n,t-1}$ removes mechanical drift from price movements, so that the measure isolates active rebalancing. We sort households into 3 turnover terciles within each of the 10 wealth brackets, which results in 30 portfolios.

2. Lottery Preference

Households with lottery preferences overweight stocks with high maximum payoffs, high idiosyncratic risk, and low prices (Kumar, 2009; Bali et al., 2011). The broader lottery-demand literature links these preferences to beta, IVOL, MAX, and skewness because probability weighting, preference for positive skewness, and salience of upside states all increase demand for lottery-like stocks (Barberis and Huang, 2008; Bali, Brown, Murray, and Tang, 2017; Bordalo, Gennaioli, and Shleifer, 2012, 2013; Boyer, Mitton, and Vorkink, 2010; Mitton and Vorkink, 2007; Cosemans and Frehen, 2021). We define four stock-level lottery signals: $\text{MAX}_{n,t}$ (maximum daily return over the past month), $\text{IVOL}_{n,t}$, $\sigma_{n,t}$ (total return volatility), and $\text{Penny}_{n,t} \in \{0,1\}$ (an indicator for low-price stocks).[8] Each signal requires at least 10 valid daily observations in its window. We map each signal to $[-1,1]$ using cross-sectional

[8] $\text{IVOL}_{n,t}$ is the standard deviation of daily residuals from a Fama-French three-factor (MKT, SMB, HML) regression estimated over a three-month rolling window. $\sigma_{n,t}$ is the standard deviation of daily returns in the prior month. $\text{Penny}_{n,t} = 1$ if the median daily price of stock n in the prior month is below 5 NOK and 0 otherwise.

ranks (denoted by tildes) and form a composite lottery score:

$$L_{n,t} = \tfrac{1}{4}\big(\widetilde{\text{MAX}}_{n,t} + \widetilde{\text{IVOL}}_{n,t} + \widetilde{\sigma}_{n,t} + \widetilde{\text{Penny}}_{n,t}\big). \tag{3}$$

Each household's lottery tilt equals the portfolio-weighted average of $L_{n,t}$:

$$m^{\text{LOT}}_{i,t} = \sum_{n} w_{i,n,t}\, L_{n,t}. \tag{4}$$

We sort on $m^{\text{LOT}}_{i,t}$ within wealth brackets to obtain 50 portfolios.

3. Disposition Effect

The disposition effect—the tendency to sell winners and hold losers—is one of the most robust household trading patterns (Odean, 1998; Calvet, Campbell, and Sodini, 2009). Define a sell indicator and a winner indicator:

$$\text{Sell}_{i,n,t} = \mathbf{1}\{Q_{i,n,t} < Q_{i,n,t-1}\}, \qquad \text{Winner}_{n,t} = \mathbf{1}\{R_{n,t-1\to t} > 0\}, \tag{5}$$

where $R_{n,t-1\to t}$ is the stock return over the preceding period. The disposition score measures how much more likely household i is to sell a stock when it is sitting on a paper gain than when it is sitting on a paper loss. A positive score means the household tends to realize gains and hold on to losses.

We estimate the disposition effect empirically by counting stock-month pairs over a rolling 24-month window. Let $N^{\text{win}}_{i,t}$ ($N^{\text{loss}}_{i,t}$) be the number of (stock, month) pairs in household i's 24-month history in which the household held a winning (losing) stock, and let $S^{\text{win}}_{i,t}$ ($S^{\text{loss}}_{i,t}$) be the subset in which the household also reduced the position. The empirical disposition score is

$$m^{\text{DISP}}_{i,t} = \frac{S^{\text{win}}_{i,t}}{N^{\text{win}}_{i,t}} - \frac{S^{\text{loss}}_{i,t}}{N^{\text{loss}}_{i,t}}. \tag{6}$$

A household holding many stocks contributes one observation per stock per month, so the score automatically aggregates across the full portfolio rather than being computed stock-by-

stock.[9] We then sort households into quintiles within wealth brackets to form 50 portfolios.

4. Extrapolation

Extrapolative investors form expectations from past returns and direct active demand toward recent winners (Barberis et al., 1998a; Greenwood and Shleifer, 2014; Barberis, Greenwood, Jin, and Shleifer, 2015; Hong and Stein, 1999; Barberis, 2018). Let $\widetilde{E}_{n,s} \in [-1, 1]$ be the cross-sectional rank of stock n's cumulative return over the 12 months ending in score month s. We remove passive portfolio drift before measuring household demand:

$$\Delta w^{\text{act}}_{i,n,s} = w_{i,n,s} - w^{\text{passive}}_{i,n,s}, \qquad \kappa_{i,s} = \sum_{n} \Delta w^{\text{act}}_{i,n,s} \widetilde{E}_{n,s}. \tag{7}$$

The monthly score $\kappa_{i,s}$ is positive when the household actively shifts portfolio weight toward recent winners and away from recent losers. Our extrapolation measure averages this sensitivity over the preceding year:

$$m^{\text{EXT}}_{i,t} = \frac{1}{|\mathcal{S}_{i,t}|} \sum_{s \in \mathcal{S}_{i,t}} \kappa_{i,s}, \qquad \mathcal{S}_{i,t} = \{s : t-12 \le s \le t-1,\ \Delta \boldsymbol{w}^{\text{act}}_{i,s} \neq \mathbf{0}\}, \tag{8}$$

with at least three active score months required. Each month, we sort eligible households on $m^{\text{EXT}}_{i,t}$ into quintiles within ten lagged stock-wealth deciles. Households that cannot be assigned remain in the residual household portfolio. As reported in Table 2, the assigned quintiles cover about 10% of stockholding households but about 33% of household stock wealth on average.

Descriptive Statistics

Table 2 summarizes the four bias measures, the number of bias-intensity bins, and score coverage. Overconfidence, lottery preference, and the disposition effect are defined for most stockholding households and most household equity wealth. Extrapolation is thinner by household count—about 10% of households—but still covers about one-third of household

[9] For example, if over the 24 months before t household i held a winning position in 22 stock-months and sold during 8 of them, and held a losing position in 20 stock-months and sold during 3 of them, then $m^{\text{DISP}}_{i,t} = 8/22 - 3/20 \approx 0.214$.

stock wealth, because the score requires repeated trading in scored stocks.

Table 2. Behavioral Bias Measures: Construction and Coverage

Bias	Bins	HH %	AUM %
Overconfidence	3	74	84
Lottery preference	5	72	82
Disposition effect	5	94	97
Extrapolation	5	10	33

This table summarizes the four behavioral bias measures. The Bins column gives the number of bias-intensity bins inside each wealth decile. Overconfidence uses three bins because portfolio turnover is heavily right-skewed in the household population; the other three measures use quintiles. HH % is the share of households with stock holdings whose score is defined in a given month. AUM % is their share of household stock wealth. Both are computed each month and averaged over the months in which the score exists. The extrapolation score requires trades in scored stocks in at least three of the past twelve months, so it covers 10% of households but 33% of household stock wealth.

Table 3 shows that the measures are only weakly related. Pairwise Spearman correlations among overconfidence, lottery preference, and the disposition effect range from 0.12 to 0.14, and correlations of those measures with extrapolation are near zero or slightly negative. The sorts therefore do not collapse to a single "biased household" type; they capture largely distinct dimensions of trading and portfolio choice.

Table 3. Pairwise Correlations of Bias Scores

	Overconfidence	Lottery	Disposition	Extrapolation
Overconfidence	1.00	0.129	0.138	−0.032
Lottery	0.129	1.00	0.115	−0.100
Disposition	0.138	0.115	1.00	−0.112
Extrapolation	−0.032	−0.100	−0.112	1.00

Each cell reports the time-series average of the cross-sectional Spearman rank correlation between two household-level bias scores. Correlations are computed each month across households for whom both scores are defined and then averaged over the sample. Pairs with extrapolation use the households with an extrapolation score, about 38,000 to 45,000 a month, against about 280,000 for the other pairs. The sample period runs from February 2006 through November 2020.

The bias ranks are persistent for three of the four measures. Table 4 reports the probability that a household remains in the same intensity quintile 12 or 36 months later. Under random reshuffling, those stay-put rates would be only 0.20. Observed diagonals are far higher for overconfidence, lottery preference, and the disposition effect: at 12 months, 0.96 of the least-active and 0.80 of the most-active overconfidence quintiles remain in place, and disposition extremes stay put with probabilities of 0.72 and 0.68. Lottery ranks are less

sticky (0.49 and 0.63) but still well above chance. Extrapolation ranks are much less persistent, with extreme-quintile stay-put rates of about 0.24 at both horizons, only modestly above chance. Even at 36 months, overconfidence, lottery preference, and the disposition effect remain above their random-assignment benchmarks.

Table 4. Transition Probabilities Across Bias Groups

Bias	Horizon	$\Pr(Q_1 \to Q_1)$	$\Pr(Q_5 \to Q_5)$	Mean diagonal
Overconfidence	12m	0.957	0.799	0.907
Overconfidence	36m	0.956	0.763	0.893
Lottery	12m	0.487	0.625	0.417
Lottery	36m	0.458	0.564	0.384
Disposition	12m	0.723	0.680	0.736
Disposition	36m	0.542	0.424	0.582
Extrapolation	12m	0.239	0.235	0.281
Extrapolation	36m	0.249	0.239	0.278

Each row reports the time-series average of transition matrices between bias quintiles at the indicated horizon. Each month, households are sorted into quintiles of the bias score across all scored households; this applies to every bias, overconfidence included. The diagonal entry $\Pr(Q_k \to Q_k)$ is the probability that a household in quintile k at month t is in quintile k at month $t + h$, among households scored at both dates. The mean diagonal averages the five diagonal entries. Under random assignment each diagonal entry equals 0.20. The sample period runs from January 2006 through November 2020.

Table 5 reports the demographic profile of the least- and most-biased households. The most-biased households tend to be younger and more often male, echoing the association between biases and demographics in the existing literature (Barber and Odean, 2001; Kumar, 2009; Calvet et al., 2009). The most lottery-biased households, for example, are on average 50 years old and 20% female, compared with 56 years old and 37% female among the least lottery-biased households. Households sorted on extrapolation are younger still, around 43 years old in both the least- and most-biased quintiles, with little gender gap across those extremes. In the raw data, only overconfidence is strongly concentrated among wealthier households: the most active tercile holds about nine times as much equity wealth as the least active tercile, whereas the lottery, disposition, and extrapolation extremes hold more similar wealth.

Table 5. Demographic Profile and Within-Group Wealth of Households by Bias

	Mean age		Share female		Wealth ratio	
Bias	least	most	least	most	uncond.	within group
Overconfidence	55.1	53.1	0.32	0.28	9.48	7.41
Lottery	55.6	50.2	0.37	0.20	0.89	0.87
Disposition	53.4	47.0	0.26	0.21	1.63	1.13
Extrapolation	43.6	43.3	0.10	0.12	1.36	1.28

The table profiles the least- and most-biased households for each bias. Each month, households are sorted into bias bins within ten groups of lagged stock wealth. The most-biased bin is the top tercile for overconfidence and the top quintile for the other three biases; the least-biased bin is the corresponding bottom tercile or quintile. For extrapolation, the bottom quintile holds the most contrarian traders, who sell after high past returns. Mean age and share female are the means within each bin, averaged across wealth groups. The wealth ratio is the stock wealth held by the most-biased bin divided by that held by the least-biased bin. It is computed unconditionally (bins formed across all households) and within wealth groups (the ratio in each group, weighted by the group's share of household equity). The sample period runs from February 2006 through November 2020.

A natural concern is that the behavioral sorts pick up alternative household characteristics rather than behavioral biases. Three tests in Appendix J.4 address this concern. First, the most-biased households underperform the least-biased for overconfidence, lottery preference, and the disposition effect. Second, because overconfidence is measured by turnover, high-turnover households might trade heavily for year-end tax reasons rather than from overconfidence. Yet, the most overconfident tercile shows no December spike in sell rates, and its December sales tilt toward winners rather than losers, the opposite of tax-motivated trading. Third, undiversified single-stock holdings (such as employer stock) could mimic lottery or high-volatility tilts without reflecting behavioral preferences. Against that alternative, portfolio concentration (maximum weight, Herfindahl index, and stock count) is flat or declining across bias bins.

2.3 Household Wealth Concentration

Because we sort behavioral intensity within wealth brackets, and because later counterfactuals ask which households move prices, the distribution of household equity wealth is a first-order input. Figure 2 compares the distribution of household equity holdings across wealth deciles for Norway and the United States, plotted on a log scale with holders in each country ranked by the size of their equity holdings. The Norwegian bars report time-series

averages of monthly inside-equity shares, using the same wealth groups that underlie the bias sorts; the U.S. bars are from the SCF 2022 using directly held stocks plus equity in trusts and managed accounts, the closest analogue to the Norwegian direct-plus-holding-company measure. Concentration is overwhelmingly a top-decile phenomenon in both countries: the bottom five deciles together hold under 1% of household equity, the ninth decile holds 4.3% in Norway and 9.1% in the United States, and the top decile holds 89.4% in Norway and 85.2% in the United States. The household sector is therefore not well summarized by the typical small retail account in either country. Most household equity dollars, and hence most potential price impact, sit in the far right tail of the holder distribution.

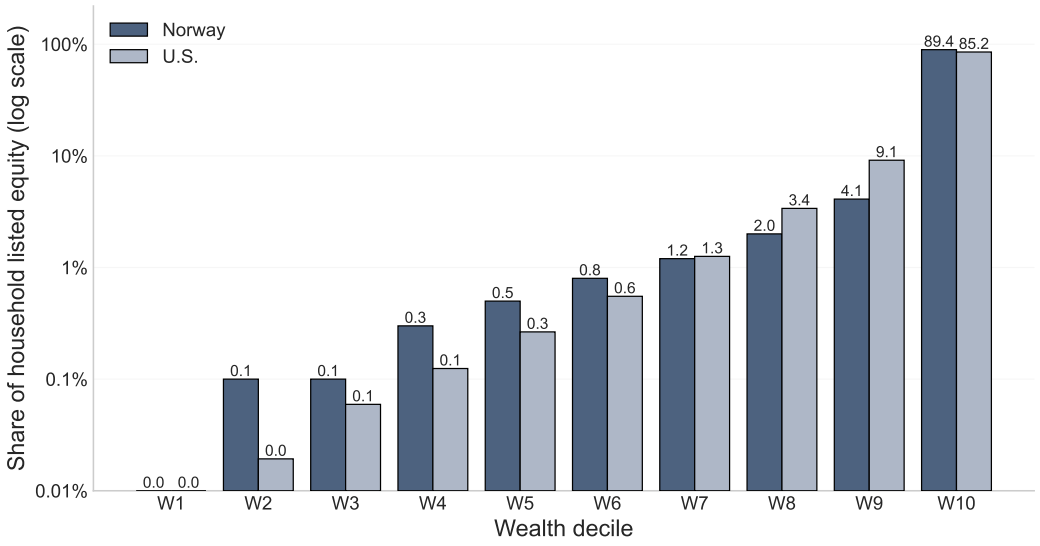


Figure 2. Concentration of household listed equity by decile, Norway versus the United States. Bars give the share of household listed equity held by each decile, on a log scale. Norway (dark) is the inside-equity share by household wealth decile, averaged monthly over 2007–2020; the United States (light) is the SCF 2022 measure of directly held stocks plus equity in trusts and managed accounts, with equity holders ranked by the size of their equity holdings. Bars below 0.01% are not shown.

2.4 Anomaly Portfolios

We construct anomaly portfolios from the Global Factor Data characteristic library of Jensen et al. (2023). Of 153 candidate characteristics, 139 have sufficient coverage in the Norwegian inside-stock panel. A characteristic must support at least 85 return months, at

least 98 months with distinct 30/70 breakpoints, and both portfolio legs in every retained month. The median retained characteristic has 119 return months. Appendix A reports every retained definition and the 14 exclusions.

Each month, we sort inside stocks using the characteristic's published orientation and 30/70 percentile breakpoints, take equal-weighted positions in the long and short legs, and rebalance monthly. All sorting variables are dated so that they are available at portfolio formation. Equal weighting is the natural specification for a broad cross-sectional test in Norway, where value weighting can place most of a portfolio on a few large firms (see, more broadly, Byun, Loudis, and Schmidt, 2026).

3 Methodology

In this section, we lay out the equilibrium system used in the counterfactual analysis. Our goal is not to estimate a fully unrestricted structural model of dynamic demand. Instead, we recover a local demand response and a local return response. Total elasticity summarizes how prices enter investor demand and, through market clearing, how wealth-flow shocks move equilibrium prices. Pass-through maps those price distortions into next-month expected returns. Appendix G gives the fuller derivation from Campbell–Shiller returns, investor demand, endogenous risk, and market clearing.

3.1 Setup and Key Objects

At month t, there are N_t inside assets. Bold lowercase symbols denote vectors, bold uppercase symbols denote matrices, and non-bold symbols denote scalars; $\mathbf{1}$ and $\boldsymbol{I}$ denote a conformable vector of ones and identity matrix. Shares outstanding are normalized to one, so prices are market equity and holdings are shares demanded. Let $\boldsymbol{p}_t, \boldsymbol{d}_t, \boldsymbol{r}_{t+1} \in \mathbb{R}^{N_t}$ denote log prices, log dividends, and one-period stock returns, and index investors by i.

For investor i, let $a_{i,t}$ denote log AUM, let $\boldsymbol{w}_{i,t} \in \mathbb{R}^{N_t}$ denote log portfolio weights, and let $\boldsymbol{q}_{i,t} \in \mathbb{R}^{N_t}$ denote log holdings, i.e. log fractions of shares outstanding demanded. Thus $a_{i,t}\mathbf{1} + \boldsymbol{w}_{i,t}$ is the log of the dollar value of investor i's demand for each stock, and subtracting

the log prices yields the following identity:

$$\boldsymbol{q}_{i,t} = a_{i,t}\mathbf{1} + \boldsymbol{w}_{i,t} - \boldsymbol{p}_t. \tag{9}$$

To map price distortions into expected returns, we need a local law of motion for prices. We write the conditional price law as

$$\mathbb{E}_t[\boldsymbol{p}_{t+1}] = \boldsymbol{c}_t + \boldsymbol{\Psi}_t \boldsymbol{p}_t, \tag{10}$$

where $\mathbb{E}_t$ denotes conditional expectation given date-t information, $\boldsymbol{c}_t$ collects time-t intercept terms that do not vary with current prices, and $\boldsymbol{\Psi}_t$ is the price-persistence matrix. If $\boldsymbol{\Psi}_t = \mathbf{0}$, a price change is expected to revert immediately next period; if $\boldsymbol{\Psi}_t = \boldsymbol{I}$, it is expected to persist one-for-one. Combining this price law with the Campbell–Shiller log-linearization gives

$$\boldsymbol{\mu}_t \equiv \mathbb{E}_t[\boldsymbol{r}_{t+1}] = \bar{\boldsymbol{\mu}}_t - \boldsymbol{\Theta}_t \boldsymbol{p}_t, \qquad \boldsymbol{\Theta}_t \equiv \boldsymbol{I} - \rho \boldsymbol{\Psi}_t, \tag{11}$$

where $\bar{\boldsymbol{\mu}}_t$ collects expected dividends and other intercept terms. The scalar ρ is the corresponding log-linearization constant: price divided by price plus dividend at the approximation point. In the baseline specification, we use $\rho = 0.9638^{1/12} \approx 0.9969$, the monthly conversion of the annual approximation point in Cochrane (2008); equivalently, $1-\rho$ is about 0.3% at the monthly horizon.

The matrix $\boldsymbol{\Theta}_t$ is the price pass-through to expected returns (Davis, 2025). In the simple one-asset case, a pass-through of 0.04 means that a 1% price drop is expected to raise next-period returns by 0.04%.

Investor demand depends on expected returns, factor risk, idiosyncratic risk, and investor-specific latent demand, meaning demand shifters not captured by expected returns or risk. Because expected returns and risk may themselves move with prices, the equilibrium price effect is governed by the local coefficient on current prices. We write the resulting reduced demand system as

$$\boldsymbol{q}_{i,t} = \boldsymbol{\kappa}_{i,t} - \boldsymbol{\eta}^{\text{tot}}_{i,t} \boldsymbol{p}_t, \tag{12}$$

where $\boldsymbol{\kappa}_{i,t}$ collects AUM, latent demand, expected dividends, and other intercept terms that

do not load on current prices. The matrix $\boldsymbol{\eta}_{i,t}^{\text{tot}}$ is the *total elasticity* matrix. It combines the expected-return channel, factor-risk channels, and the idiosyncratic-risk channel into the single price coefficient needed for market clearing. We use this total object in the main counterfactuals because it is exactly what maps wealth-flow shocks into equilibrium prices. As some have argued, the total elasticity may not be an invariant structural primitive between: pass-through and risk responses may differ across shocks (Davis et al., 2026; He et al., 2025). The concern is that the counterfactual analysis may differ from the estimated shocks, presenting a Lucas critique style issue. To address this Lucas critique, Appendix M therefore recovers the expected-return and risk components and varies the pass-through and endogenous-risk mappings, and we show that the results are robust to these variations.

3.2 Equilibrium Flow Multiplier

We now close the model in equilibrium. Following Davis et al. (2026); van der Beck (2022); van der Beck et al. (2025), we embed the reduced demand system in log-linearized market clearing around observed ownership shares. The multiplier below is a result of combining demand with market clearing, not an additional assumption. The log-linearized market-clearing condition is

$$\mathbf{0} = \sum_i \bar{\boldsymbol{S}}_{i,t} \boldsymbol{q}_{i,t}, \tag{13}$$

where $\bar{\boldsymbol{S}}_{i,t}$ is the diagonal matrix of investor i's observed ownership shares, the point around which we linearize.[10]

Throughout, we distinguish observed time differences from counterfactual perturbations. For any object x_t, let $\Delta^{\text{obs}} x_t \equiv x_{t+1} - x_t$ denote the realized change from month t to month $t+1$, and let $\Delta^{\text{cf}} x_t \equiv x_t^{\text{cf}} - x_t$ denote the counterfactual change relative to the observed date-t

[10] Without the log-linearization, market clearing is $\log\left(\sum_i \exp(\boldsymbol{q}_{i,t})\right) = \mathbf{0}$. The left-hand side is the log of aggregate demand. The right-hand side is the log of supply: because shares outstanding are normalized to one stock-by-stock, supply is $\mathbf{1}$ in levels and $\mathbf{0}$ in logs. A first-order expansion around observed ownership shares yields (13). van der Beck (2022) show that this log-linearized market-clearing condition is a very good approximation even with larger deviations from observed ownership shares.

equilibrium. For counterfactual log-AUM shifts $\{\Delta^{\text{cf}} a_{i,t}\}_i$, define the induced flow vector

$$\Delta^{\text{cf}} \boldsymbol{f}_t \equiv \sum_i \bar{\boldsymbol{S}}_{i,t} \Delta^{\text{cf}} a_{i,t} \mathbf{1}. \tag{14}$$

The vector $\Delta^{\text{cf}} \boldsymbol{f}_t$ represents the stock-level fraction of shares outstanding associated with the inflows and outflows.

PROPOSITION 1 (Flow Multiplier)**.** *Under the price law, reduced-demand representation, and log-linearized market-clearing system above, if investors experience counterfactual log-AUM shifts* $\{\Delta^{\text{cf}} a_{i,t}\}_i$ *while* $\boldsymbol{\eta}_{i,t}^{\text{tot}}$ *is held fixed, then the equilibrium price response to the induced flow vector* $\Delta^{\text{cf}} \boldsymbol{f}_t$ *satisfies*

$$\Delta^{\text{cf}} \boldsymbol{p}_t = \underbrace{\left(\sum_i \bar{\boldsymbol{S}}_{i,t} \boldsymbol{\eta}_{i,t}^{\text{tot}} \right)^{-1}}_{\equiv \boldsymbol{\mathcal{M}}_t} \Delta^{\text{cf}} \boldsymbol{f}_t. \tag{15}$$

The inverse of the aggregate total-elasticity matrix is the price multiplier for counterfactual wealth flows.

The proof is in Appendix H. Proposition 1 is the key equilibrium mapping for our analysis: once we estimate total elasticities, observed ownership shares determine how counterfactual wealth flows translate into prices.

3.3 Characteristics, Factors, and Loadings

Holdings and returns at the stock level are unbalanced panels, and so we use a factor model to govern spillovers as suggested by Haddad et al. (2025). Consider four characteristics: CAPM beta, book equity, ESG greenness, and ROE. Each month, we rank-transform these into $[-0.5, 0.5]$ and orthonormalize the resulting columns by right-whitening. We then do Kelly, Pruitt, and Su (2019) instrumented principal component analysis (IPCA) with these orthonormal characteristics, which Appendix I.1 shows is equivalent to doing PCA on the characteristic-weighted portfolios. We do not reduce the number of characteristics from this process, so we end up with $K = 4$ characteristics, which are the original characteristics that are just rotated and scaled. Let $\boldsymbol{X}_t \in \mathbb{R}^{N_t \times K}$ denote the resulting orthonormal

basis, and let $\boldsymbol{x}'_{n,t}$ be row n. This last PCA step ensures these characteristics are also factor weights as well as factor loadings, so we can simply write $\boldsymbol{\beta}_t = \boldsymbol{X}_t$. Importantly, these resulting characteristics are still orthonormal, so $\boldsymbol{X}'_t\boldsymbol{X}_t = \boldsymbol{I}$. Appendix I.1 gives additional details. This is the same procedure used, for example, by Chaudhry and Davis (2026).

3.4 Factor Model Elasticity and Pass-Through

We use the low-rank functional form suggested by Haddad et al. (2025) for the elasticity matrix and the pass-through matrix. Following Koijen et al. (2022), we pool the elasticity coefficients within investor group. Let $\mathcal{G}(i)$ denote investor i's group, and use $\mathcal{G}$ to denote a generic group. The demand regressions are estimated on investor-stock-month observations; the group index only determines which elasticity coefficients are shared across investors. We write

$$\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t} = (1-\eta_{0,\mathcal{G}})\boldsymbol{I} + \boldsymbol{X}_t\boldsymbol{\Gamma}_{\mathcal{G}}\boldsymbol{X}'_t, \qquad \boldsymbol{\Psi}_t = \psi_0\boldsymbol{I} + \boldsymbol{X}_t\boldsymbol{\Gamma}_{\psi}\boldsymbol{X}'_t, \tag{16}$$

where $\eta_{0,\mathcal{G}}$ and ψ_0 are scalar own-stock coefficients, and $\boldsymbol{\Gamma}_{\mathcal{G}}$ and $\boldsymbol{\Gamma}_{\psi}$ are $K \times K$ matrices that capture spillovers through the common characteristic portfolios.

These restrictions have a simple investor-level regression interpretation. Let

$$\boldsymbol{g}_t \equiv \boldsymbol{p}_t - \boldsymbol{b}_t \tag{17}$$

denote the month-t valuation gap, where $\boldsymbol{b}_t$ is log book equity. Let $x_{n,k,t}$ denote the (n,k) entry of $\boldsymbol{X}_t$, let $\Gamma_{\mathcal{G},kl}$ denote the (k,l) entry of $\boldsymbol{\Gamma}_{\mathcal{G}}$, and define the factor-portfolio valuation gap

$$g^{\text{fac}}_{l,t} \equiv \sum_{m=1}^{N_t} x_{m,l,t}g_{m,t}. \tag{18}$$

Then the elasticity specification implies the investor-stock-month panel regression:

$$w_{i,n,t} = \alpha_{i,t} + \sum_{k=1}^{K} x_{n,k,t}\beta_{k,i} + \eta_{0,\mathcal{G}(i)}g_{n,t} - \sum_{k=1}^{K}\sum_{l=1}^{K} x_{n,k,t}\Gamma_{\mathcal{G}(i),kl}\, g^{\text{fac}}_{l,t} + u_{i,n,t}, \tag{19}$$

To estimate pass-through at the horizon over which valuation distortions resolve, let $R^{(12)}_{n,t\to t+12}$ denote stock n's realized compounded return from the end of month t through the end of

month $t+12$. We estimate the forecasting regression

$$R^{(12)}_{n,t\to t+12} = \alpha_n + \gamma_t + \boldsymbol{c}'_{n,t}\boldsymbol{\delta} + \beta^{(12)}_0 g_{n,t} + \sum_{k=1}^{K}\sum_{l=1}^{K} x_{n,k,t}\Gamma^{(12)}_{R,kl}\, g^{\text{fac}}_{l,t} + \nu^{(12)}_{n,t}, \tag{20}$$

where $\alpha_{i,t}$ is an investor-month effect, $\beta_{k,i}$ is investor i's slope on characteristic k, α_n and γ_t are stock and month fixed effects, and $\boldsymbol{c}_{n,t}$ contains the four date-t characteristics (book equity, CAPM beta, ESG greenness, and ROE). We convert the 12-month return coefficients to monthly pass-through units by defining

$$\theta_0 = -\frac{\beta^{(12)}_0}{12}, \qquad \boldsymbol{\Gamma}_\Theta = -\frac{\boldsymbol{\Gamma}^{(12)}_R}{12}, \qquad \boldsymbol{\Theta}_t = \theta_0 \boldsymbol{I} + \boldsymbol{X}_t \boldsymbol{\Gamma}_\Theta \boldsymbol{X}'_t, \tag{21}$$

and recover the corresponding monthly price-law parameters from $\boldsymbol{\Psi}_t = (\boldsymbol{I} - \boldsymbol{\Theta}_t)/\rho$. Thus the return horizon is genuinely twelve months; the division by 12 expresses the estimated response in the monthly units used by the counterfactuals. The coefficients $\eta_{0,\mathcal{G}(i)}$ and $\Gamma_{\mathcal{G}(i),kl}$ are common within investor group, so investors in the same group share the same own-stock and spillover elasticities even though the absorbed intercepts and characteristic slopes are investor specific. In other words, the left-hand side remains $w_{i,n,t}$ throughout; what is pooled is the coefficient block indexed by $\mathcal{G}(i)$. The same investor-level left-hand side is used in the broad demand system and in each of the four bias systems, with $\mathcal{G}(i)$ changing only through the grouping rule. The direct and double-sum terms capture own-stock responses and spillovers through the common factor portfolios: the shock first aggregates into $g^{\text{fac}}_{l,t}$, and then maps back into stock space according to each stock's loading $x_{n,k,t}$.

There is a second way to motivate this functional form for the elasticity matrix. Haddad et al. (2025) point out that the elasticity matrix with mean-variance demand is the inverse of the covariance matrix, Σ_t^{-1}.[11] Consider a model where the idiosyncratic variance, σ^2, is uncorrelated and the same across stocks. If $\boldsymbol{\Omega}$ is the covariance matrix of factors, then we can write

$$\boldsymbol{\Sigma}_t = \sigma^2 \boldsymbol{I} + \boldsymbol{X}_t \boldsymbol{\Omega} \boldsymbol{X}'_t,$$

[11] This is actually the Jacobian of demand with respect to prices, but we adopt the Haddad et al. (2025) parlance that this is the natural-units elasticity in this model.

Note that here we are using the fact that $\boldsymbol{X}_t = \boldsymbol{\beta}_t$. Orthonormality and the Woodbury identity give

$$\boldsymbol{\Sigma}_t^{-1} = \underbrace{\sigma^{-2}}_{\text{analogue of } 1-\eta_{0,\mathcal{G}}} \boldsymbol{I} \quad + \boldsymbol{X}_t \underbrace{\left[-\sigma^{-4}\left(\boldsymbol{\Omega}^{-1} + \sigma^{-2}\boldsymbol{I}_K\right)^{-1}\right]}_{\text{analogue of } \boldsymbol{\Gamma}_{\mathcal{G}}} \boldsymbol{X}_t'.$$

So $\boldsymbol{\Sigma}_t^{-1}$ is the same form as $\boldsymbol{\eta}_{\mathcal{G},t}^{tot}$ above in (16).

Finally, these regressions require excluded instruments for both the direct price terms and the spillover terms. In the investor-level demand panel, if $\hat{g}_{i,n,t}$ is a valid instrument for $g_{n,t}$, define the instrumented factor-portfolio gap by $\hat{g}_{i,l,t}^{\text{fac}} \equiv \sum_{m=1}^{N_t} x_{m,l,t}\hat{g}_{i,m,t}$. Then the interactions $x_{n,k,t}\hat{g}_{i,l,t}^{\text{fac}}$ are natural instruments for the spillover regressors. In the pass-through panel, the corresponding stock-level instruments are $\hat{g}_{n,t}$ and $\hat{g}_{l,t}^{\text{fac}} \equiv \sum_{m=1}^{N_t} x_{m,l,t}\hat{g}_{m,t}$. Their validity requires the usual exclusion restriction: conditional on the included fixed effects and controls, they can affect holdings demand and future prices only through the endogenous valuation-gap regressors. Appendix I.2 states this construction explicitly.

3.5 Estimation

We estimate the demand system in (19) and the pass-through system in (20) with instrumental variables (IV). Following the sticky-mandate instrument of Koijen and Yogo (2019) and the grouped implementation in Koijen et al. (2022), we first define the stock-level instrumented log market-equity measure

$$IV_{n,t} \equiv \log\left(\sum_{j\in\mathcal{J}_t(n)} \frac{A_{j,t}}{1+N_{j,t}}\right), \tag{22}$$

where $\mathcal{J}_t(n)$ is the set of institutional accounts whose investment universe contains stock n in month t, $A_{j,t}$ is account j's AUM, and $N_{j,t}$ is the number of stocks in that mandate. In the underlying instrument construction, the universe is built from the previous 36 months of holdings and restricted to accounts with at least 95% overlap between current holdings and that historical universe.

For the investor-level demand panel, we use the leave-out-own-account analogue:

$$IV_{i,n,t} \equiv \log\left(\sum_{j\in\mathcal{J}_t(n)\setminus\{i\}} \frac{A_{j,t}}{1+N_{j,t}}\right). \tag{23}$$

If investor i is not one of the institutional accounts in $\mathcal{J}_t(n)$, then this coincides with $IV_{n,t}$.

We then form the stock-level and investor-level instrumented valuation gaps

$$\hat{g}_{n,t} \equiv IV_{n,t} - b_{n,t}. \tag{24}$$

$$\hat{g}_{i,n,t} \equiv IV_{i,n,t} - b_{n,t}. \tag{25}$$

The corresponding interacted instruments are built from $\hat{g}^{\text{fac}}_{i,l,t} = \sum_{m=1}^{N_t} x_{m,l,t}\hat{g}_{i,m,t}$ on the demand side and $\hat{g}^{\text{fac}}_{l,t} = \sum_{m=1}^{N_t} x_{m,l,t}\hat{g}_{m,t}$ on the pass-through side. Thus the demand system uses the investor-specific leave-out-own instrument $\hat{g}_{i,n,t}$ and its interacted analogues, while the pass-through system uses the stock-level instrument $\hat{g}_{n,t}$ and its interacted analogues. On the demand side, we residualize the panel with respect to investor-month effects and investor-specific characteristic slopes, and then estimate investor-level 2SLS with coefficients constrained to be common within investor group. On the pass-through side, we estimate the analogous stock-month IV panel with stock and month fixed effects, using the realized compounded 12-month return as the outcome and mapping its coefficients to monthly units through (21). One may naturally be concerned that the K^2 spillover terms in (19) and (20) could potentially overfit the data. Therefore, we use a standard empirical-Bayes shrinkage term to discount any potential noise, as discussed in Appendix I.3. Appendix K shows that the three main results are also similar when the low-rank bilinear spillover block is replaced with a local-similarity kernel specification. Appendix L reaches similar conclusions when the same kernel system is estimated month by month and then averaged over time, again using the same empirical-Bayes logic to shrink noisy monthly spillover coefficients before forming the time-averaged system. We stress that the results are similar with and without the shrinkage, but our main specification uses this shrinkage to ease any concerns about overly noisy estimates.

Operationally, we estimate two families of demand systems. First, we estimate the broad grouped system for all investors, with the seven groups in $\mathcal{K}$ from Section 2; this is the system used for the household-versus-nonhousehold variance decomposition. Second, for each household bias family b, we re-estimate the investor-level demand system after replacing the single household block with the reported household bias buckets for that family, while leaving the nonhousehold groups unchanged. Let $\mathcal{G}_b$ denote the resulting group set for bias family b, and let $\mathcal{G}_{b,\ell} \subseteq \mathcal{G}_b$ denote the household sleeves that belong to reported bin ℓ. This second step is necessary because the across-bias variance decompositions and expected-return attributions ask which household bias buckets are the relevant marginal demanders. To answer that question, the demand elasticities must be allowed to differ across those buckets rather than being pooled into one household coefficient block. The stock-level pass-through system is estimated once and then combined with the broad or bias-specific demand system depending on the counterfactual exercise.

All panel standard errors are two-way clustered by stock and month. For the nonlinear variance-decomposition and anomaly-attribution objects, the reported table standard errors come from 5,000 joint bootstrap draws that sample the first-step coefficients from their estimated covariance matrices and resample months in common 12-month circular blocks. Each draw is shared across tilts, anomalies, designs, and pooled contrasts, and inference retains finite, stable systems that can be solved directly.

3.6 Counterfactual Experiments

The payoff from the estimated system is that it delivers two counterfactual objects directly.

3.6.1 Variance Decomposition

We consider a variance decomposition similar in spirit to Koijen and Yogo (2019), but where the ordering of terms in the variance decomposition does not affect the results. The variance exercise begins from observed month-to-month changes in group ownership shares on the common stock universe present in both month t and month $t+1$. Let $\mathcal{K}$ denote the

seven investor groups. For investor group $\mathcal{G} \in \mathcal{K}$, let $\boldsymbol{s}_{\mathcal{G},t}$ and $\boldsymbol{s}_{\mathcal{G},t+1}$ denote the group share vectors on that common universe, and let $\Delta^{\text{obs}}\boldsymbol{p}_t \equiv \boldsymbol{p}_{t+1} - \boldsymbol{p}_t$ denote the realized log market-equity change. The group's latent flow shock is the residual demand shift that rationalizes the observed change in holdings given the estimated total elasticity:

$$\hat{\boldsymbol{z}}_{\mathcal{G},t} \equiv \log \boldsymbol{s}_{\mathcal{G},t+1} \;-\; \log \boldsymbol{s}_{\mathcal{G},t} \;+\; \hat{\boldsymbol{\eta}}^{\text{tot}}_{\mathcal{G},t} \Delta^{\text{obs}}\boldsymbol{p}_t. \tag{26}$$

Multiplying by the initial share vector converts this log-demand shift into stock-level flow pressure,

$$\hat{\boldsymbol{f}}_{\mathcal{G},t} \equiv \boldsymbol{s}_{\mathcal{G},t} \odot \hat{\boldsymbol{z}}_{\mathcal{G},t}, \tag{27}$$

where $\odot$ denotes the elementwise product. For any subset of investor groups $\mathcal{A} \subseteq \mathcal{K}$, the model-implied price vector is

$$\widehat{\Delta^{\text{obs}}\boldsymbol{p}}_t(\mathcal{A}) \equiv \hat{\boldsymbol{\mathcal{M}}}_t \sum_{\mathcal{G} \in \mathcal{A}} \hat{\boldsymbol{f}}_{\mathcal{G},t}. \tag{28}$$

For any subset of groups $\mathcal{A}$, define equal-weighted and value-weighted variance in the usual way:

$$V_t^{\text{EW}}(\mathcal{A}) \equiv \text{Var}_n\left(\widehat{\Delta^{\text{obs}}p}_{n,t}(\mathcal{A})\right), \qquad V_t^{\text{VW}}(\mathcal{A}) \equiv \sum_n \omega_{n,t+1}\left(\widehat{\Delta^{\text{obs}}p}_{n,t}(\mathcal{A}) - \Delta^{\bar{\text{obs}}}p_t^{\omega}(\mathcal{A})\right)^2, \tag{29}$$

where $\omega_{n,t+1}$ are next-month market-cap weights normalized to sum to one. We report two nonnegative allocations of these monthly objects and a signed Shapley allocation that includes covariance between investor contributions.

The first is the *Standalone-variance* decomposition. For each player k, define the standalone price effect

$$\widehat{\Delta^{\text{obs}}\boldsymbol{p}}^{\text{solo}}_{k,t} \equiv \hat{\boldsymbol{\mathcal{M}}}_t \hat{\boldsymbol{f}}_{k,t},$$

and allocate variance by

$$s^{\text{stand}}_{k,t}(V) \equiv \frac{V_t\left(\widehat{\Delta^{\text{obs}}\boldsymbol{p}}^{\text{solo}}_{k,t}\right)}{\sum_{h \in \mathcal{K}} V_t\left(\widehat{\Delta^{\text{obs}}\boldsymbol{p}}^{\text{solo}}_{h,t}\right)}. \tag{30}$$

The second is an R^2 *Shapley* decomposition. Let

$$y_t \equiv \widehat{\Delta^{\text{obs}} \boldsymbol{p}}_t(\mathcal{K}) = \sum_{k \in \mathcal{K}} \widehat{\Delta^{\text{obs}} \boldsymbol{p}}^{\text{solo}}_{k,t} \tag{31}$$

denote the full model-implied price change. For any coalition $\mathcal{A} \subseteq \mathcal{K}$, we ask how well the full vector y_t can be approximated using only the standalone price-impact vectors of the groups in $\mathcal{A}$. Let $\tilde{\omega}_{n,t} = 1$ for the equal-weighted decomposition and $\tilde{\omega}_{n,t} = \omega_{n,t+1}$ for the value-weighted decomposition. Define the fitted coefficients by the weighted least-squares problem

$$(a_t(\mathcal{A}), \{\beta_{k,t}(\mathcal{A})\}_{k \in \mathcal{A}}) \equiv \arg \min_{a, \{\beta_k\}_{k \in \mathcal{A}}} \sum_n \tilde{\omega}_{n,t} \left(y_{n,t} - a - \sum_{k \in \mathcal{A}} \beta_k \widehat{\Delta^{\text{obs}} p}^{\text{solo}}_{k,n,t} \right)^2, \tag{32}$$

and let

$$\widehat{y}_{n,t}(\mathcal{A}) \equiv a_t(\mathcal{A}) + \sum_{k \in \mathcal{A}} \beta_{k,t}(\mathcal{A}) \widehat{\Delta^{\text{obs}} p}^{\text{solo}}_{k,n,t}. \tag{33}$$

The coalition value is then the weighted in-sample fit

$$v_t(\mathcal{A}) \equiv 1 - \frac{\sum_n \tilde{\omega}_{n,t} \left(y_{n,t} - \widehat{y}_{n,t}(\mathcal{A}) \right)^2}{\sum_n \tilde{\omega}_{n,t} \left(y_{n,t} - \bar{y}_t^{\tilde{\omega}} \right)^2}, \tag{34}$$

where $\bar{y}_t^{\tilde{\omega}}$ is the $\tilde{\omega}$-weighted mean of y_t. Equivalently, $v_t(\mathcal{A})$ is the weighted R_t^2 from regressing the full model-implied price change on the standalone components in coalition $\mathcal{A}$.

We then assign group k its Shapley contribution to fit:

$$s_{k,t}^{R^2} \equiv \sum_{\mathcal{A} \subseteq \mathcal{K} \setminus \{k\}} \frac{|\mathcal{A}|! \left(|\mathcal{K}| - |\mathcal{A}| - 1 \right)!}{|\mathcal{K}|!} \left[v_t(\mathcal{A} \cup \{k\}) - v_t(\mathcal{A}) \right]. \tag{35}$$

This differs from the Standalone-variance decomposition. The Standalone-variance object asks how much variance each standalone vector has on its own. The R^2 Shapley object instead asks how much each group's standalone price-impact pattern helps span the *full* cross-section of model-implied price changes after the other groups' standalone patterns are already in the regression. A group therefore receives a large R^2 Shapley share when it explains parts of y_t that the other standalone components cannot already replicate well.

Because adding regressors cannot reduce R^2, these marginal contributions are nonnegative. Since y_t lies in the span of all standalone components by construction, $v_t(\mathcal{K}) = 1$, so the R^2 Shapley shares sum to one.

The third is a *signed Shapley* decomposition, which allocates the variance of the sum of investor contributions, including their covariances. Write $x_{k,t} \equiv \widehat{\Delta^{\text{obs}} \boldsymbol{p}}^{\text{solo}}_{k,t}$ for group k's stock-level price contribution, so that $y_t = \sum_{k \in \mathcal{K}} x_{k,t}$. Let $\text{Var}_{w,t}$ and $\text{Cov}_{w,t}$ denote cross-sectional variance and covariance using either equal or value stock weights. For a coalition $\mathcal{A} \subseteq \mathcal{K}$, define its value as $u_t(\mathcal{A}) = \text{Var}_{w,t}(\sum_{k \in \mathcal{A}} x_{k,t})$, holding the full-market multiplier, stock sample, and weights fixed. Averaging the marginal contribution over all group orderings assigns group k its own variance plus one copy of each covariance with another group:

$$\begin{aligned} \phi_{k,t} &= \text{Var}_{w,t}(x_{k,t}) + \sum_{h \neq k} \text{Cov}_{w,t}(x_{k,t}, x_{h,t}) = \text{Cov}_{w,t}(x_{k,t}, y_t), \\ s^{\text{signed}}_{k,t} &= \frac{\phi_{k,t}}{\text{Var}_{w,t}(y_t)}, \qquad \sum_{k \in \mathcal{K}} s^{\text{signed}}_{k,t} = 1. \end{aligned} \tag{36}$$

Each pair's covariance enters total variance twice and is split equally between the two groups. A negative share identifies a contribution that offsets the full price-change pattern; shares can exceed one when other contributions offset it. Shares are defined when total variance is positive and are reported as percentages after averaging across months.

The household-bias decomposition applies the same three methods to hypothetical shocks to the assigned household bins. For each bias family b, let $\mathcal{B}$ denote those bins and let $\hat{\boldsymbol{f}}^{50}_{b,\ell,t} = -0.5 \sum_{g \in \mathcal{G}_{b,\ell}} \boldsymbol{s}_{g,t}$ denote the stock-level flow pressure from a hypothetical withdrawal of 50% of bin ℓ's AUM. Shocks to residual households and nonhousehold groups are zero; their demand responses still enter the full-market multiplier. Let

$$\hat{\boldsymbol{H}}_{b,t} \equiv \sum_{g \in \mathcal{G}_b} \bar{\boldsymbol{S}}_{g,t} \hat{\boldsymbol{\eta}}^{\text{tot}}_{g,t}, \qquad \hat{\boldsymbol{\mathcal{M}}}_{b,t} \equiv \hat{\boldsymbol{H}}^{-1}_{b,t}, \tag{37}$$

denote the aggregate elasticity matrix and multiplier, including all investor groups. Each matrix row indexes the stock whose demand changes and each column indexes the stock whose log price changes; the inverse maps stock-level flow pressure into log-price changes.

For any coalition $\mathcal{A} \subseteq \mathcal{B}$, define

$$\Delta \boldsymbol{p}_{b,t}^{50}(\mathcal{A}) \equiv \hat{\boldsymbol{\mathcal{M}}}_{b,t} \sum_{\ell \in \mathcal{A}} \hat{\boldsymbol{f}}_{b,\ell,t}^{50}. \tag{38}$$

We apply the three allocations to these bin contributions, with total variance defined as the variance of $\Delta \boldsymbol{p}_{b,t}^{50}(\mathcal{B})$. Thus the bin shares sum to one within the assigned-household experiment. They measure the price-variance effects of the hypothetical shocks, rather than a decomposition of observed return variance. Each bias classification is analyzed separately because households can belong to more than one classification.

3.6.2 Expected-Return Attribution

The main expected-return experiment is a within-bias dollar-neutral reallocation. For the disposition effect, lottery preference, and extrapolation, the donor set $\mathcal{D}_b$ contains every $Q4$ and $Q5$ wealth-by-tilt sleeve and the receiver set $\mathcal{R}_b$ contains every $Q1$ and $Q2$ sleeve. For overconfidence, $\mathcal{D}_b$ contains the $T3$ sleeves and $\mathcal{R}_b$ contains the $T1$ sleeves. Let $A_{b,t}^D$ and $A_{b,t}^R$ denote their respective month-t assets under management. Each donor position is reduced by 50%, and the same total dollars are added to all receiver positions through a common proportional increase:

$$\Delta^{\mathrm{cf}} a_{g,t}^b \equiv \begin{cases} -\lambda, & g \in \mathcal{D}_b, \\ \lambda A_{b,t}^D / A_{b,t}^R, & g \in \mathcal{R}_b, \\ 0, & \text{otherwise}, \end{cases} \qquad \lambda = 0.5. \tag{39}$$

The $Q3$ and $T2$ sleeves, unassigned and residual households, and every nonhousehold group are unchanged. The construction is dollar neutral in every month. The choice $\lambda = 0.5$ is purely a normalization. Under the linearized system, the induced price distortion and the expected-return contribution scale proportionally with λ, so a different percentage shock would just rescale the reported entries one-for-one. The implied stock-level flow is

$$\Delta^{\mathrm{cf}} \boldsymbol{f}_t^b \equiv \sum_g \boldsymbol{s}_{g,t} \Delta^{\mathrm{cf}} a_{g,t}^b, \tag{40}$$

and the corresponding price distortion is

$$\Delta^{\text{cf}}\boldsymbol{p}_t^b \equiv \hat{\boldsymbol{\mathcal{M}}}_{b,t}\Delta^{\text{cf}}\boldsymbol{f}_t^b. \tag{41}$$

For extrapolation, $\boldsymbol{s}_{g,t}$ is the month-t portfolio held by households grouped according to the lagged active-demand score in (8). The counterfactual therefore reallocates capital between the portfolios of households with different histories of extrapolation. It does not reverse the trades used to construct the score. We then map this price distortion into next-month expected returns through the baseline one-step pass-through matrix. At the stock level, the actual effect of implementing the counterfactual—the removal effect—is

$$\Delta^{\text{cf}}\boldsymbol{\mu}_t^b \equiv \boldsymbol{\mu}_t^{\text{cf},b} - \boldsymbol{\mu}_t^{\text{obs}} = -\hat{\boldsymbol{\Theta}}_t\Delta^{\text{cf}}\boldsymbol{p}_t^b. \tag{42}$$

For anomaly portfolio P with weights $\boldsymbol{W}_t^P$ from Section 2.4, we report the negative of this removal effect: the observed-minus-counterfactual contribution

$$\mathcal{C}_{b,t}^P \equiv -12\,(\boldsymbol{W}_t^P)'\Delta^{\text{cf}}\boldsymbol{\mu}_t^b = 12\,(\boldsymbol{W}_t^P)'\hat{\boldsymbol{\Theta}}_t\Delta^{\text{cf}}\boldsymbol{p}_t^b. \tag{43}$$

Let $\bar{m}_b$ denote the time-series average fraction of total household AUM moved by the counterfactual for bias family b. To compare the four counterfactuals at the same scale, we report

$$\mathcal{C}_{b,P}^{20} \equiv \frac{0.20}{\bar{m}_b}\mathcal{C}_{b,P}. \tag{44}$$

The four counterfactuals move 23.3%, 19.8%, 32.7%, and 7.9% of average household AUM for the disposition effect, lottery preference, overconfidence, and extrapolation, respectively. These numbers differ from 20% for a simple reason: sorting households into quintiles does not necessarily mean that each quintile has the same amount of capital. Fortunately, since the effects are (log) linear, we can simply rescale them. Equation (44) changes only the units: groups, relative shocks, dollar neutrality, elasticities, pass-through, anomaly assignments, and realized returns remain fixed. The 20% target applies on average over the sample

months, not in every month.[12] Thus $\Delta^{\text{cf}}\boldsymbol{\mu}_t^b$ answers how expected returns would change under the reallocation, whereas $\mathcal{C}_{b,t}^P$ measures the corresponding contribution in the observed economy. The portfolio weights collapse the stock-level contribution to an anomaly spread, and the factor of 12 annualizes it. This decomposition follows the logic of Davis et al. (2025a). Appendix N examines dynamic decay when anomaly portfolios are held fixed or reconstituted over time; the conclusions remain unchanged.

The reported cell for bias family b and anomaly P is the time-series mean of $\mathcal{C}_{b,t}^P$, not the removal effect $\Delta^{\text{cf}}\mu$, across months in which the anomaly spread is defined; the tables report joint-bootstrap standard errors below the estimates. Positive entries mean that the behavioral tilt raises the published long-minus-short return, while negative entries mean that it lowers that return.

3.6.3 Predictions for Expected-Return Attribution

A rich empirical and theoretical literature links behavioral biases to stock-return anomalies. Before examining the counterfactual estimates, we collect the directional predictions that literature articulates for the biases we measure. For example, for overconfidence, Daniel et al. (1998) and Daniel, Hirshleifer, and Subrahmanyam (2001) argue that biased self-attribution and overconfident beliefs generate under- and overreaction that show up as value, momentum, and related cross-sectional premia, while Scheinkman and Xiong (2003) emphasize speculative overpricing when disagreement meets short-sale constraints. For prospect-theory and lottery demand, Barberis and Huang (2008), Kumar (2009), Bali et al. (2011), and Barberis et al. (2021) link probability weighting and related preferences to low subsequent returns on volatile, skewed, and issuance-intensive stocks. For the disposition effect, Grinblatt and Han (2005) and Frazzini (2006) predict momentum and post-news underreaction from the tendency to sell winners and hold losers. For extrapolation, Lakonishok et al. (1994) predict value premia when investors extrapolate past growth too far.

We map each prediction to the anomaly portfolios that fall under the proposed mechanism, then group those portfolios into economic categories for interpretation. For example,

[12] Appendix F reports the same contribution evidence without the 20% AUM normalization, and the results are quite similar.

overconfidence's value prediction is tested with the ten fundamental-to-price characteristics in our value category: assets-to-market, book-to-market equity and enterprise value, dividend yield, EBITDA-to-enterprise value, free cash flow-to-price, intrinsic value-to-market, earnings-to-price, operating cash flow-to-market, and sales-to-market. Similarly, the disposition effect's momentum prediction is tested with the six price-momentum and 52-week-high characteristics in our momentum category: returns from $t-3$ to $t-1$, $t-6$ to $t-1$, $t-9$ to $t-1$, $t-12$ to $t-1$, and $t-12$ to $t-7$, together with the ratio of current price to the prior-year high. Analogous mappings cover market beta, duration/growth, volatility, investment and issuance, post-news drift, turnover, maximum returns, distress and safety, and skewness; Appendix C lists every characteristic in each category, and Appendix B records the economic argument and source for every bias–anomaly link we test. In total, this yields 18 bias–category predictions spanning 51 distinct anomalies.

4 Results

4.1 Elasticity and Pass-Through Estimates

We begin with the demand and pass-through estimates that underpin all three counterfactual exercises. The two tables below summarize the key coefficients and compare them with existing estimates from the literature. The headline numbers are consistent with the demand-system literature: own-price elasticities are modest and economically plausible across all investor groups, cross-stock spillovers are small in both the demand and pass-through systems, and the implied price pass-through to expected returns falls squarely within the range of prior estimates.

Table 6 reports the demand estimates. Panel A shows the direct own-stock term $\eta_{0,\mathcal{G}}$ together with a compact summary of the spillover block $\Gamma_{\mathcal{G}}$. The household own-stock coefficient is about 0.09, while the institutional coefficient is about 0.21. The diagonal entries of $\Gamma_{\mathcal{G}}$ are modest and the average off-diagonal entry remains small across all broad investor groups, even when the most positive and most negative off-diagonal terms are reported explicitly.

Panel A: Group-level coefficient summaries

Group	$\eta_{0,\mathcal{G}}$	Γ_{11}	Γ_{22}	Γ_{33}	Γ_{44}	Avg |Off|	Avg Off	Min Off	Max Off
Households	0.0917	-0.0079	0.0328	0.0291	0.0632	0.0328	0.0089	-0.0606	0.0923
	(0.1239)	(0.0566)	(0.0390)	(0.0484)	(0.0283)	(0.0147)	(0.0146)	(0.0391)	(0.0569)
Institutions	0.2077	0.0004	0.0392	-0.0181	0.0719	0.0319	0.0069	-0.0545	0.0818
	(0.2674)	(0.0447)	(0.0319)	(0.0568)	(0.0376)	(0.0115)	(0.0180)	(0.0390)	(0.0340)
Governments	-1.2431	-0.0134	-0.0074	-0.0547	0.0102	0.0106	0.0033	-0.0161	0.0265
	(0.7957)	(0.0169)	(0.0128)	(0.0195)	(0.0085)	(0.0066)	(0.0064)	(0.0201)	(0.0205)
Listed firms	0.6791	-0.0370	0.0110	0.0361	-0.0122	0.0157	0.0114	-0.0181	0.0312
	(0.2353)	(0.0190)	(0.0107)	(0.0288)	(0.0138)	(0.0056)	(0.0055)	(0.0231)	(0.0185)
Foreign	0.6758	-0.0149	0.0486	0.0917	0.0875	0.0429	-0.0140	-0.0809	0.0808
	(0.1399)	(0.0466)	(0.0249)	(0.0546)	(0.0372)	(0.0136)	(0.0138)	(0.0318)	(0.0335)
Other	0.5282	0.0301	0.0234	0.0017	0.0093	0.0196	-0.0001	-0.0421	0.0747
	(0.2040)	(0.0320)	(0.0174)	(0.0218)	(0.0164)	(0.0076)	(0.0088)	(0.0236)	(0.0339)
Non-listed firms	0.7082	0.0081	0.0307	0.0568	0.0577	0.0326	0.0132	-0.0462	0.1123
	(0.1471)	(0.0397)	(0.0226)	(0.0335)	(0.0219)	(0.0084)	(0.0110)	(0.0190)	(0.0505)

Panel B: Implied stock-level elasticity averages

Group	Avg diag($\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$)	Avg |off-diag|($\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$)	Avg off-diag($\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$)
Households	0.9093	0.0010	-0.0000081
	(0.1233)	(0.0003)	(0.0000071)
Institutions	0.7931	0.0010	-0.0000065
	(0.2667)	(0.0002)	(0.0000076)
Governments	2.2426	0.0005	0.0000045
	(0.7954)	(0.0002)	(0.0000029)
Listed firms	0.3208	0.0005	0.0000001
	(0.2352)	(0.0001)	(0.0000030)
Foreign	0.3260	0.0013	-0.0000147
	(0.1394)	(0.0003)	(0.0000061)
Other	0.4723	0.0007	-0.0000045
	(0.2038)	(0.0002)	(0.0000038)
Non-listed firms	0.2930	0.0011	-0.0000106
	(0.1469)	(0.0002)	(0.0000045)

Table 6. Demand estimates by broad investor group. Panel A reports the direct own-stock coefficient and summaries of the low-rank spillover matrix for each broad investor group. Off-diagonal summaries use the 12 off-diagonal entries of $\Gamma_{\mathcal{G}}$ and report both signed and absolute averages. Panel B reports the corresponding stock-level average diagonal together with the signed and absolute average off-diagonal entries of $\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t} = (1 - \eta_{0,\mathcal{G}})\boldsymbol{I} + \boldsymbol{X}_t \Gamma_{\mathcal{G}} \boldsymbol{X}'_t$, averaged across all stocks and months. Standard errors are reported below the estimates. For the min/max rows, the standard error corresponds to the reported extremum.

Panel A: Matrix coefficient summaries

Matrix ψ_0 or θ_0	Γ_{11} (×1000)	Γ_{22} (×1000)	Γ_{33} (×1000)	Γ_{44} (×1000)	Avg |Off| (×1000)	Avg Off (×1000)	Min Off (×1000)	Max Off (×1000)
$\boldsymbol{\Psi}_t$	0.9808	-22.410	-9.083	-30.179	18.734	17.465	1.446	-37.186	27.905
	(0.0078)	(14.950)	(10.639)	(15.099)	(15.054)	(8.614)	(3.346)	(14.343)	(14.998)
$\boldsymbol{\Theta}_t$	0.0222	22.342	9.055	30.086	-18.677	17.411	-1.442	-27.819	37.072
	(0.0078)	(14.904)	(10.606)	(15.053)	(15.008)	(8.588)	(3.336)	(14.952)	(14.299)

Panel B: Implied stock-level matrix averages

Matrix	Avg diag	Avg |off-diag| (×1000)	Avg off-diag (×1000)
$\boldsymbol{\Psi}_t$	0.9805	0.4561	0.002973
	(0.0079)	(0.1705)	(0.001845)
$\boldsymbol{\Theta}_t$	0.0225	0.4547	-0.002964
	(0.0079)	(0.1700)	(0.001840)

Table 7. Pass-through estimates in $\boldsymbol{\Psi}_t$ and $\boldsymbol{\Theta}_t$ space. We estimate the predictive response of the realized compounded 12-month return to the current valuation gap and its low-rank interactions, then divide the return-response block by -12 to express pass-through in monthly units and recover $\boldsymbol{\Psi}_t = (\boldsymbol{I} - \boldsymbol{\Theta}_t)/\rho$. Panel A reports the own-price coefficients together with summaries of the low-rank blocks for the monthly price-law matrix and the implied pass-through matrix. Off-diagonal summaries use the 12 off-diagonal entries of the corresponding Γ block and report both signed and absolute averages. Panel B reports the corresponding stock-level average diagonal together with the signed and absolute average off-diagonal entries of $\boldsymbol{\Psi}_t = \psi_0 \boldsymbol{I} + \boldsymbol{X}_t \Gamma_\psi \boldsymbol{X}_t'$ and $\boldsymbol{\Theta}_t = \theta_0 \boldsymbol{I} + \boldsymbol{X}_t \Gamma_\theta \boldsymbol{X}_t'$, averaged across all stocks and months. Standard errors are two-way clustered by stock and month and are reported below the estimates. For the min/max rows, the standard error corresponds to the reported extremum.

Panel B translates those coefficient estimates back into the implied stock-level elasticity matrix $\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$. The average diagonal entry is about 0.91 for households and about 0.79 for institutions, while the average off-diagonal entry is economically tiny for every group. The estimates therefore preserve some structured substitution through the factor block without generating large diffuse cross-stock spillovers in the stock-level matrix. This is in keeping with Chaudhry and Davis (2026), who measure stock-level cross-elasticities in a large variety of ways and find uniformly that cross-elasticities are economically tiny.

Table 7 reports the pass-through estimates directly in $\boldsymbol{\Psi}_t$ and $\boldsymbol{\Theta}_t$ space. The own-stock coefficient in the 12-month return regression is -0.266, which maps through (21) to a monthly own pass-through θ_0 of 0.022 (standard error 0.008). Panel B maps the estimates back into the stock-level matrices: the average diagonal of $\boldsymbol{\Psi}_t$ is about 0.98, the average diagonal of $\boldsymbol{\Theta}_t$ is about 0.023, and the average signed off-diagonal terms in both matrices are economically tiny. The estimated $\boldsymbol{\Theta}_t$ has positive eigenvalues in every sample month, and no stability correction is required. The magnitude is well in line with the literature. Davis (2025) estimate a price pass-through of about 0.01 over a monthly horizon. Koijen et al. (2022) have an annual pass-through of about 0.1, yielding a monthly pass-through of about 0.008. Davis et al. (2025b) estimates a higher pass-through of 0.048 per month, and thus our pass-through estimate sits within the range of the existing literature.

4.2 Households Account for a Large Share of Return Variance

In our first counterfactual exercise, we ask how much of the cross-sectional variation in stock returns can be attributed to household demand. The question is related to a long debate about whether investor heterogeneity generates excess price variation. Theoretical work on noise traders establishes that investors who act on signals unrelated to fundamentals can survive in equilibrium and move prices (De Long, Shleifer, Summers, and Waldmann, 1990), but the aggregate price impact of any particular investor class depends on how much capital it commands and how distinct its portfolio tilts are across stocks. Whether the household sector satisfies these conditions has been difficult to assess without comprehensive investor-level data and a structural model.

Panel A: Variance-share decomposition						
Group	R^2 Shapley		Standalone		Signed Shapley	
	EW	VW	EW	VW	EW	VW
Households	33.4	22.0	34.4	18.5	36.1	21.9
	(2.8)	(2.0)	(3.6)	(2.5)	(3.3)	(2.5)
Institutions	25.4	28.9	23.7	29.7	26.4	31.4
	(3.4)	(4.2)	(3.7)	(5.3)	(4.0)	(5.9)
Foreign	11.6	11.7	11.7	11.8	9.3	9.2
	(1.5)	(1.6)	(2.0)	(2.2)	(2.0)	(2.4)
Governments	2.4	13.9	2.4	15.5	2.4	14.4
	(0.6)	(1.9)	(0.9)	(3.4)	(0.7)	(3.2)
Listed firms	7.7	5.6	7.4	5.7	6.9	5.2
	(1.4)	(1.1)	(1.6)	(1.2)	(1.8)	(1.4)
Other	9.6	9.2	11.0	10.4	10.3	10.3
	(1.8)	(1.5)	(2.1)	(1.9)	(2.1)	(2.0)
Non-listed firms	10.1	8.6	9.4	8.5	8.5	7.6
	(1.5)	(1.3)	(2.1)	(2.0)	(2.0)	(2.0)

Panel B: Variance share relative to mean % AUM						
Group	R^2 Shapley		Standalone		Signed Shapley	
	EW	VW	EW	VW	EW	VW
Households	1.86	1.23	1.92	1.03	2.01	1.22
	(0.15)	(0.11)	(0.20)	(0.14)	(0.18)	(0.14)
Institutions	0.66	0.75	0.61	0.77	0.68	0.81
	(0.09)	(0.11)	(0.10)	(0.14)	(0.10)	(0.15)
Foreign	0.73	0.74	0.74	0.75	0.59	0.58
	(0.10)	(0.10)	(0.12)	(0.14)	(0.12)	(0.15)
Governments	0.11	0.63	0.11	0.71	0.11	0.65
	(0.03)	(0.09)	(0.04)	(0.15)	(0.03)	(0.15)

Table 8. Household and nonhousehold variance shares. Panel A allocates model-implied cross-sectional price-change variance under three decompositions and equal (EW) or next-month market-capitalization (VW) stock weights. Historical demand shocks are recovered between consecutive calendar months. Signed shares allocate covariance with the sum of all seven investor contributions and can be negative. Panel B divides each share by its mean inside-asset share for the four largest AUM groups. All methods use February 2007–March 2020 (158 transitions). Standard errors appear below the estimates. All three methods are recomputed in 4,999 accepted joint coefficient draws with common 12-month circular blocks of months and fixed fitted shrinkage weights; AUM shares are held fixed.

Table 8 reports the three variance decompositions. Panel A gives the raw variance-share allocations, while Panel B rescales those allocations by each group's mean inside-asset share for the four largest AUM groups. Households own about 18% of inside assets, but their

modeled variance share is substantially larger under all three methods. Under our preferred R^2 Shapley decomposition, households account for 33.4% of equal-weighted and 22.0% of value-weighted modeled variance. Under the Standalone-variance decomposition, the corresponding shares are 34.4% and 18.5%. The signed Shapley allocation gives households 36.1% of equal-weighted and 21.9% of value-weighted modeled variance, confirming the same broad pattern when covariance is allocated directly.

The broad conclusion is therefore robust to the decomposition method: households account for far more cross-sectional return variation than their ownership share would suggest. In Panel B, the household contribution is about 1.9 times its AUM share in equal-weighted terms and about 1.2 times its AUM share in value-weighted terms, while institutions and foreign investors remain below one on that normalized scale. The result carries two implications. First, the household sector cannot be treated as price-irrelevant noise in equilibrium models of the cross-section. Second, isolating the behavioral origins of household demand, the goal of the next two exercises, is a quantitatively meaningful question: understanding which households drive this disproportionate variance channel has direct implications for how behavioral biases enter equilibrium prices.

4.3 Biased, Wealthy Households Drive the Household Variance Channel

In this section, we ask whether the household variance channel is driven by biased households. We decompose the variance generated by hypothetical shocks to assigned household bins within each bias measure and ask how much of it is accounted for by the most biased households. The decomposition has two further dimensions. First, we split the bias channel across the wealth distribution. Because the top household wealth decile holds almost 90% of household AUM, whether the variance channel is concentrated in the top decile or spread across the distribution has direct implications for which households are the price-relevant behavioral agents. Second, we ask how low-dimensional the top-tail bias tilts are across wealth deciles. This speaks to whether biased households across the distribution hold a small number of common tilted portfolios or whether the tilts are largely sleeve-specific.

Bias	R^2 Shapley		Standalone		Signed Shapley	
	EW	VW	EW	VW	EW	VW
Overconfidence	69.9	66.0	75.9	81.9	70.6	73.4
	(4.7)	(2.9)	(4.4)	(3.9)	(4.7)	(3.4)
Lottery preference	89.8	70.8	86.4	69.4	93.5	71.6
	(2.9)	(4.8)	(3.1)	(4.8)	(3.1)	(5.0)
Disposition effect	61.2	49.5	62.3	51.6	61.6	51.8
	(6.7)	(4.4)	(6.7)	(6.5)	(6.8)	(4.8)
Extrapolation	46.7	48.6	46.9	51.9	46.9	49.4
	(6.8)	(5.5)	(6.8)	(6.6)	(6.8)	(5.8)

Table 9. Fraction of the variance generated by hypothetical shocks to assigned household bins accounted for by the most biased households. Each bin receives a 50% AUM withdrawal; residual-household and nonhousehold shocks are zero, and all investor elasticities enter the multiplier. Overconfidence reports the top tercile; the other measures report the top two quintiles. EW and VW denote equal and next-month market-capitalization stock weights. Signed shares use covariance with the sum of assigned-bin contributions, divided by the variance of that sum. Standard errors appear below the estimates. All three methods use February 2007–December 2016 (119 months), and are jointly recomputed in 4,999 accepted coefficient draws with common 12-month circular blocks and fixed fitted shrinkage weights.

Table 9 reports the share of the variance generated by the assigned household bins accounted for by the most biased households. For overconfidence, the top tercile accounts for 69.9% of the preferred equal-weighted channel. For lottery preference, the disposition effect, and extrapolation, the top two quintiles account for 89.8%, 61.2%, and 46.7%. For the first three measures, the households with the strongest behavioral bias therefore account for most of this hypothetical variance channel. The corresponding signed Shapley shares are 70.6%, 93.5%, 61.6%, and 46.9%, respectively. For extrapolation, the signed equal-weighted share is imprecisely estimated relative to its share of assigned household AUM.

Figure 3 compares the signed variance allocation with the capital controlled by the most biased households. We divide the top group's assigned-bin variance share by its share of assigned household AUM, then multiply by the broad household sector's signed variance-to-AUM ratio. This scaled comparison combines the historical broad-sector result with the hypothetical bin experiment; it is not a direct full-market allocation from one demand system.

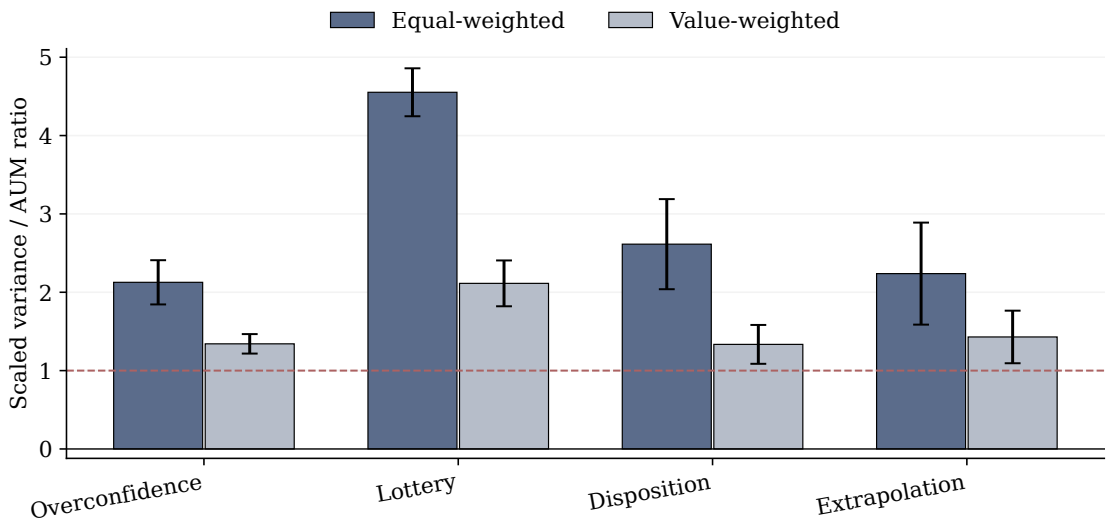


Figure 3. Scaled signed variance-to-AUM comparison for the most biased households. The top group's signed share of assigned-bin variance is divided by its share of assigned household AUM and multiplied by the broad household sector's signed variance-to-AUM ratio. The broad sector uses historical shocks; the bins use hypothetical 50% AUM shocks. The multiplier in each exercise includes all investors. The dashed line marks one. Brackets show ±2 standard errors, propagating uncertainty in the within-household allocation while holding AUM shares and the broad-sector scaling at their respective point estimates. The top group is the highest overconfidence tercile or the highest two quintiles for the other measures.

In Figure 4, we further decompose the variance contribution of the most biased households by wealth. Within each bias measure, we allocate that contribution between households in the top wealth decile and those in the bottom nine deciles, holding the denominator fixed at the variance generated by the assigned bins for the measure. The bottom-nine bar therefore measures that segment's contribution to the aggregate channel, not concentration relative to its much smaller wealth base.

Under the signed Shapley decomposition, strongly biased households in the top wealth decile account for 58–85% of the equal-weighted assigned-bin variance for overconfidence, lottery preference, and the disposition effect. The remaining nine deciles together account for only 3–8%. The wealth gradient extends to extrapolation: the top-decile sleeves contribute 44.6% of its assigned-bin variance, compared with 2.2% for the bottom nine deciles. The price-relevant behavioral household is therefore overwhelmingly a wealthy household.

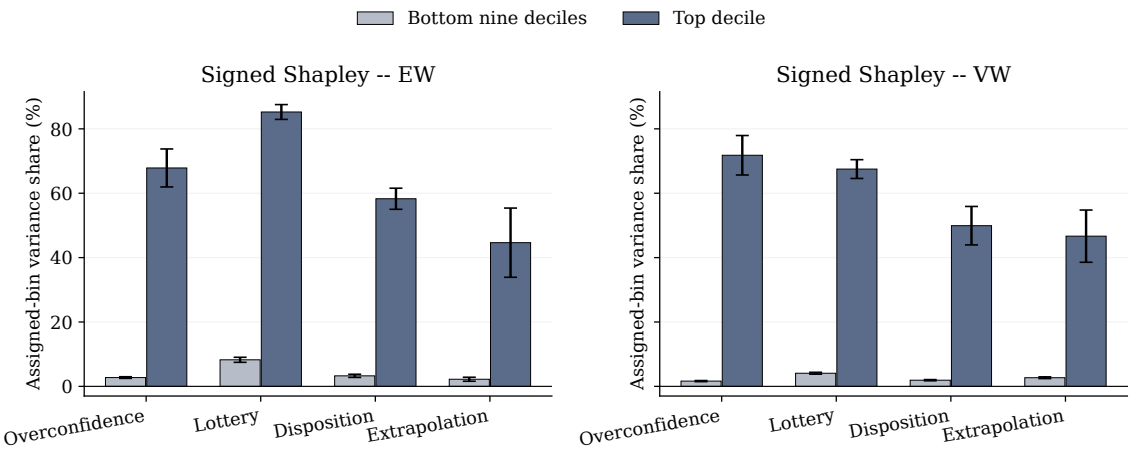


Figure 4. Signed Shapley contributions of the most biased households by wealth segment. The top wealth decile and the bottom nine deciles are allocated shares of the same assigned-bin variance target. Bin contributions use hypothetical 50% AUM shocks and the multiplier includes all investors. The most biased households are the top overconfidence tercile and the top two quintiles for the other measures. EW and VW denote equal and value stock weights. Brackets show ±2 standard errors from 999 circular 12-month block draws, holding fitted parameters fixed. These brackets measure month-to-month sampling variation and exclude parameter-estimation uncertainty. Wealth contributions add to their corresponding signed bin shares in Table 9.

These results raise the question of how correlated behavioral portfolio tilts are across the wealth distribution. If the most biased households in different wealth segments hold similar active positions, their demand shocks reinforce one another; if the tilts are largely idiosyncratic, wealth concentration alone drives the price impact. Figure 5 answers this by stacking the active tilts of the most biased households within each wealth decile and summarizing the shared variation with principal components. Across the four bias measures, the first two uncentered principal components explain 58–75% of active-tilt variation. Biased households across the wealth distribution therefore share a small number of common portfolio directions.

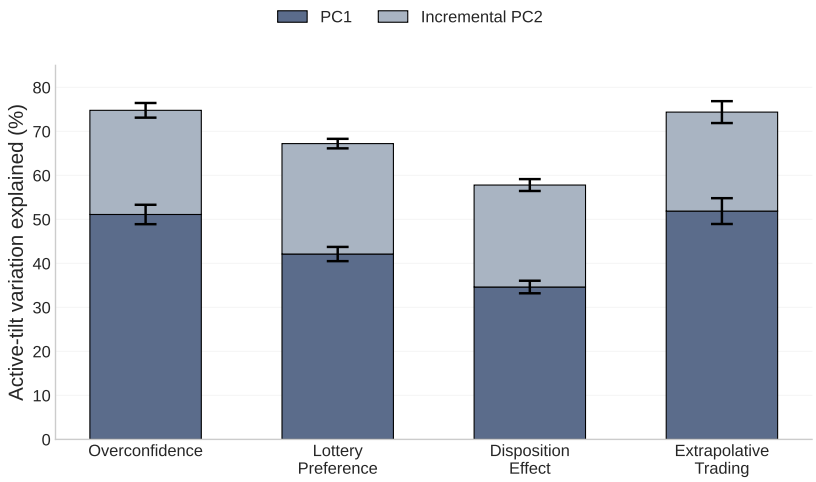


Figure 5. First two PCA components of most-biased behavioral tilts. For each bias measure and month, we form most-biased household sleeves within wealth deciles, stack their stock-level active tilts, and compute the share of equal-sleeve active-tilt variation explained by uncentered principal components. Bars report time-series averages: the dark segment is the first principal component and the light segment is the incremental contribution of the second. Error bars report ±2 Newey–West standard errors for the first-PC share and for the cumulative share explained by the first two components. The most-biased sleeves are the highest tercile for overconfidence, giving 10 sleeves, and the top two quintiles for lottery preference, the disposition effect, and extrapolation, giving 20 sleeves. Each active tilt is measured relative to all non-residual households in the same wealth decile.

4.4 Biased Households Contribute to Anomaly Returns

In our second counterfactual exercise, we shift from the second moment to the first. We implement the within-bias reallocation defined in Section 3.6.2 and trace the observed-minus-counterfactual contribution to next-period expected returns. As defined in (43), the reported sign is $\mathcal{C} = \mu^{\text{obs}} - \mu^{\text{cf}} = \boldsymbol{\Theta}\Delta^{\text{cf}}\boldsymbol{p}$: positive estimates raise the published long-minus-short return and negative estimates lower it.

We start by reporting the average contribution across anomalies for each bias measure in Table 10. For each bias measure, the table reports the average contribution across the predicted anomalies. The last row then reports the average across all bias measures. The average contribution is positive for each bias measure, implying that the anomaly spreads would be lower if household demand were shifted from the most to the least biased investors. The contribution is strongest for lottery preference (325.0 basis points per year), followed by

Table 10. Pooled tests of creator predictions (common 20% AUM scale)

Behavioral tilt	Positive sign	Average (bp)	Bootstrap SE	One-sided p
Disposition effect	6/9	70.0	40.5	0.042
Lottery preference	25/30	325.0	116.2	0.003
Overconfidence	36/42	73.0	25.5	0.002
Extrapolation	6/10	54.0	41.5	0.097
Equal tilt average	73/91	130.5	44.3	0.002

Notes: The table reports annualized equal-weighted contributions for the creator predictions. Cells are equally weighted within each tilt, and the headline equally weights the four tilt averages. Positive estimates agree with the prediction. Standard errors use exactly 5,000 common joint bootstrap draws; one-sided p-values use the standard-normal distribution.

overconfidence (73.0), the disposition effect (70.0), and extrapolation (54.0). Averaging over the bias measures gives a contribution of 130 basis points per year. As a point of reference, this corresponds to 27% of the mean realized equal-weight anomaly spread over our sample period.[13]

Figure 6 reports the average contribution for each bias–anomaly-category pair. Consistent with lottery and prospect-theory preferences generating low subsequent returns on volatile, skewed, and issuance-intensive stocks (Barberis and Huang, 2008; Kumar, 2009; Bali et al., 2011; Barberis et al., 2021), lottery-preferring households make large contributions to volatility, maximum-return, issuance, and distress-and-safety spreads (about 390–620 basis points per year); they also contribute positively to value and skewness. Consistent with overconfidence and speculative disagreement producing cross-sectional mispricing (Daniel et al., 1998, 2001; Scheinkman and Xiong, 2003), overconfident households contribute most to market-beta portfolios (about 200 basis points) and also raise duration/growth, volatility, and investment spreads. Consistent with the disposition effect generating momentum (Grinblatt and Han, 2005; Frazzini, 2006), disposition-prone households contribute about 115 basis points to momentum spreads. Extrapolation contributes positively, but more modestly, to its predicted value spreads (Lakonishok et al., 1994).

Overall, the contributions are positive for 16 of the 18 predictions, and 10 remain significant at the 5% level after Benjamini–Hochberg adjustment for multiple testing. Because the

[13] Appendix E reports the complete multiple-testing audit.

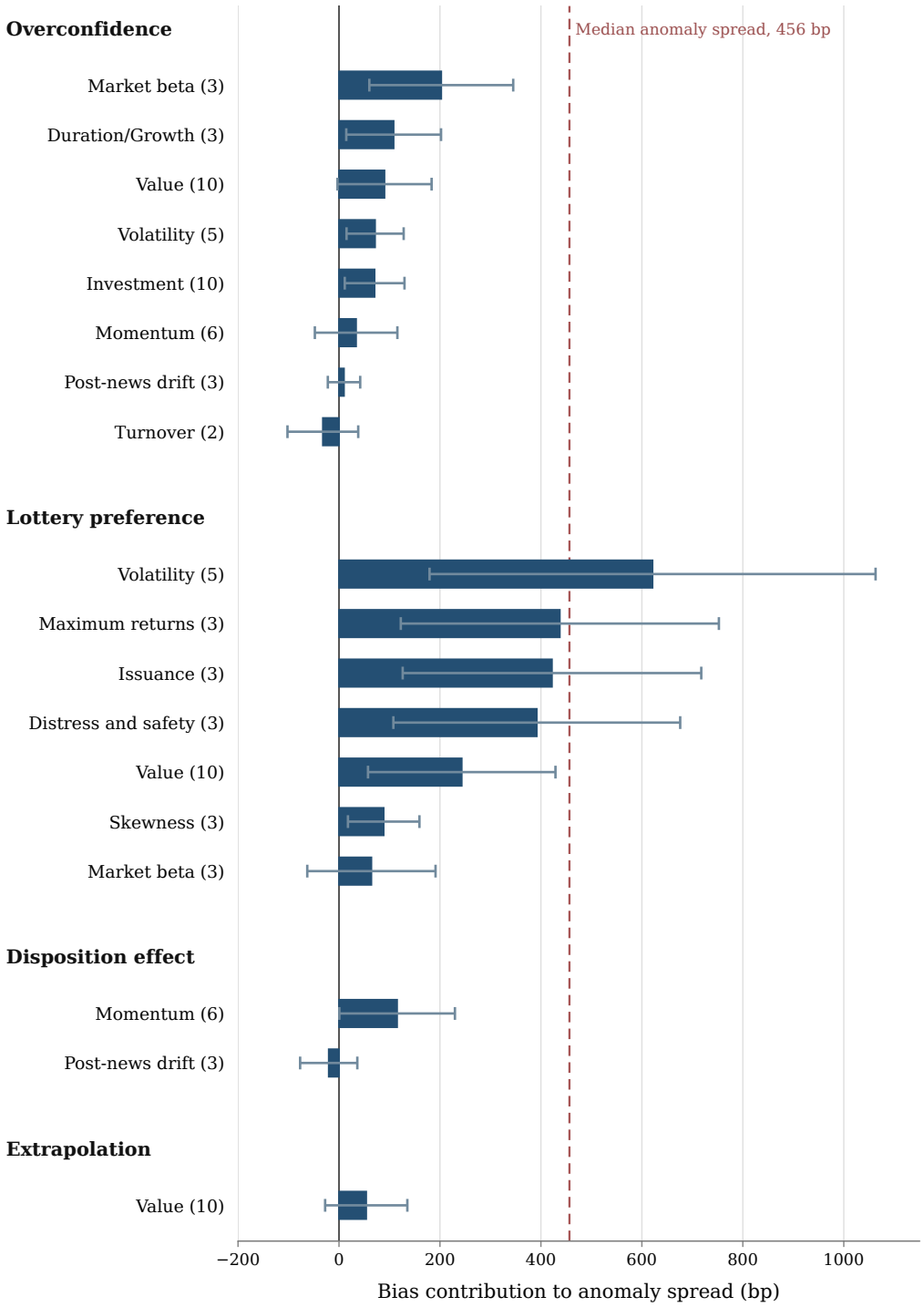


Figure 6. Anomaly contributions from biased households. Each bar is the simple average of the annualized equal-weighted contributions to the published long-minus-short spreads in the category on the common 20% AUM scale; positive values raise the spread and negative values lower it. Parentheses give the number of anomalies. Whiskers report 95% normal-bootstrap confidence intervals from 5,000 common joint draws. The dashed line is the median realized spread among the 51 unique anomalies represented in the figure. Appendix C lists the category membership and reports estimates, standard errors, confidence intervals, and global Benjamini–Hochberg q-values for all 18 bars.

counterfactual reallocates only 20% of household wealth from the most to the least biased investors, these category contributions are meaningful relative to the scale of the demand shift.[14]

Together, the two counterfactual exercises paint a consistent picture. Households move the cross-section by substantially more than their ownership share would suggest. The most biased and wealthy households drive much of that variance channel. Across a broad anomaly universe, biased households create expected returns in the directions predicted by behavioral mechanisms. Behavioral demand therefore survives aggregation and shapes both equilibrium price variation and expected returns.

4.5 Placebo Tests

We end with placebo tests that apply the same expected-return calculation to anomaly portfolios outside the literature-predicted set for each behavioral tilt. Within each tilt, we equally weight these anomaly portfolios and recompute their average contribution inside each of the 5,000 common joint bootstrap draws. The overall placebo equally weights the four tilt-level estimates.

Table 11. Placebo tests using anomalies outside the literature predictions

Placebo	Anomaly cells	Estimate (bp)	Bootstrap SE	95% CI	BH q
Overall placebo	465	-10.4	7.2	[-24.6, 3.8]	0.311
Overconfidence	97	16.0	12.1	[-7.7, 39.7]	0.311
Lottery preference	109	-9.4	15.2	[-39.1, 20.3]	0.667
Disposition effect	130	2.8	12.9	[-22.4, 28.1]	0.827
Extrapolation	129	-50.9	24.3	[-98.6, -3.3]	0.180

Notes: The table reports annualized equal-weighted expected-return contributions on the common 20% AUM scale. Within each behavioral tilt, the placebo equally weights all anomaly portfolios outside that tilt's set of literature predictions. The overall placebo equally weights the four tilt-level estimates. Standard errors use exactly 5,000 common joint bootstrap draws. The final column reports Benjamini–Hochberg q-values computed jointly across the five rows from two-sided normal p-values.

Table 11 shows that the placebo estimates are statistically insignificant after adjustment for multiple comparisons. The overall estimate is -10.4 basis points per year, with a 95%

[14] Appendix C reports the exact category estimates and membership. Appendix E reports the associated FDR adjustments.

confidence interval from −24.6 to 3.8 basis points and a Benjamini–Hochberg q-value of 0.311. The four tilt-level estimates range from −50.9 to 16.0 basis points per year, and every adjusted q-value exceeds 0.10. The contribution estimates therefore concentrate in the anomaly portfolios linked to each behavioral mechanism.

5 Conclusion

Household behavioral biases matter quantitatively for asset prices. Households own less than one-fifth of the market but account for up to one-third of cross-sectional return variance. Within that channel, the economically important component comes from the most biased households, especially those in the top wealth decile. Debiasing households by relocating 20% of household wealth from the most- to the least-biased investors compresses anomaly spreads by 130 basis points a year, or 27% of the mean realized spread. Across 18 bias–category predictions spanning 51 anomalies, 16 have the expected positive sign and 10 remain significant at the 5% level after multiple-testing adjustment.

These findings speak to the debate over rational risk compensation versus behavioral mispricing. While most economists now agree that biased investors can affect prices due to limits to arbitrage, how important biased investors are for stock returns remains debated. Returns alone cannot separate behavioral mispricing from alternative explanations (Kozak et al., 2018). By measuring behavioral biases in investor portfolios and tracing them through an equilibrium demand system, we quantify how important household biases are for both cross-sectional price variation and expected anomaly returns.

References

Ang, Andrew, and Joseph Chen, 2002, Asymmetric correlations of equity portfolios, *Journal of Financial Economics* 63, 443–494.

Baker, Malcolm, Jeremy C. Stein, and Jeffrey Wurgler, 2003, When does the market matter? stock prices and the investment of equity-dependent firms, *Quarterly Journal of Economics* 118, 969–1005.

Baker, Malcolm, and Jeffrey Wurgler, 2002, Market timing and capital structure, *Journal of Finance* 57, 1–32.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645–1680.

Balasubramaniam, Vimal, John Y Campbell, Tarun Ramadorai, and Benjamin Ranish, 2023, Who owns what? a factor model for direct stockholding, *The Journal of Finance* 78, 1545–1591.

Bali, Turan G., Stephen J. Brown, Scott Murray, and Yi Tang, 2017, A lottery-demand-based explanation of the beta anomaly, *Journal of Financial and Quantitative Analysis* 52, 2369–2397.

Bali, Turan G, Nusret Cakici, and Robert F Whitelaw, 2011, Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of financial economics* 99, 427–446.

Barber, Brad M, and Terrance Odean, 2001, Boys Will Be Boys Gender Overconfidence and Common Stock Investment, *Quarterly Journal of Economics* 116, 261–292.

Barber, Brad M., and Terrance Odean, 2013, The behavior of individual investors, volume 2 of *Handbook of the Economics of Finance*, 1533–1570 (Elsevier).

Barber, Brad M, Terrance Odean, and Ning Zhu, 2008, Do retail trades move markets?, *The Review of Financial Studies* 22, 151–186.

Barberis, Nicholas, 2018, Psychology-based models of asset prices and trading volume, in *Handbook of behavioral economics: applications and foundations 1*, volume 1, 79–175 (Elsevier).

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2015, X-CAPM: An extrapolative capital asset pricing model, *Journal of Financial Economics* 115, 1–24.

Barberis, Nicholas, and Ming Huang, 2008, Stocks as lotteries: The implications of probability weighting for security prices, *American Economic Review* 98, 2066–2100.

Barberis, Nicholas, Lawrence J Jin, and Baolian Wang, 2021, Prospect theory and stock market anomalies, *The Journal of Finance* 76, 2639–2687.

Barberis, Nicholas, Abhiroop Mukherjee, and Baolian Wang, 2016, Prospect theory and stock returns: An empirical test, *The review of financial studies* 29, 3068–3107.

Barberis, Nicholas, and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 161–199.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998a, A model of investor sentiment, *Journal of Financial Economics* 49.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998b, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler, 2005, Comovement, *Journal of financial economics* 75, 283–317.

Barberis, Nicholas, and Richard Thaler, 2003, A survey of behavioral finance, in George M. Constantinides, Milton Harris, and René M. Stulz, eds., *Handbook of the Economics of Finance*, volume 1, 1053–1128 (Elsevier).

Barberis, Nicholas, and Wei Xiong, 2009, What drives the disposition effect? an analysis of a long-standing preference-based explanation, *The Journal of Finance* 64, 751–784.

Benjamini, Yoav, and Yosef Hochberg, 1995, Controlling the false discovery rate: A practical and powerful approach to multiple testing, *Journal of the Royal Statistical Society: Series B (Methodological)* 57, 289–300.

Betermier, Sebastien, Laurent E. Calvet, Samuli Samuli Knüpfer, and Jens S. Kvaerner, 2024, Investor factors, *The Journal of Finance* forthcoming.

Betermier, Sebastien, Laurent E. Calvet, and Paolo Sodini, 2017, Who are the value and growth investors?, *Journal of Finance* 72, 5–46.

Bhamra, Harjoat S, and Raman Uppal, 2019, Does household finance matter? small financial errors with large social costs, *American Economic Review* 109, 1116–1154.

Black, Fischer, 1976, Studies of stock price volatility changes, in *Proceedings of the 1976 Meetings of the Business and Economic Statistics Section*, 177–181 (American Statistical Association).

Boehmer, Ekkehart, Charles M Jones, Xiaoyan Zhang, and Xinran Zhang, 2021, Tracking retail investor activity, *Journal of Finance* 76, 2249–2305.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2019, Diagnostic expectations and stock returns, *Journal of Finance* 74, 2839–2874.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2012, Salience theory of choice under risk, *Quarterly Journal of Economics* 127, 1243–1285.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2013, Salience and asset prices, *American Economic Review Papers and Proceedings* 103, 623–628.

Bouchaud, Jean-Philippe, Philipp Krueger, Augustin Landier, and David Thesmar, 2019, Sticky expectations and the profitability anomaly, *Journal of Finance* 74, 639–674.

Boyer, Brian, Todd Mitton, and Keith Vorkink, 2010, Expected idiosyncratic skewness, *Review of Financial Studies* 23, 169–202.

Byun, Sung Je, Johnathan Loudis, and Lawrence D. W. Schmidt, 2026, A tale of two market returns: The broad market factor and the idiosyncratic financial factor, Working Paper 35243, National Bureau of Economic Research.

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

Calvet, Laurent E, John Y Campbell, and Paolo Sodini, 2009, Measuring the financial sophistication of households, *American economic review* 99, 393–398.

Campbell, John Y., 2006, Household finance, *Journal of Finance* 61, 1553–1604.

Campbell, John Y., and Tarun Ramadorai, 2025, Household finance in retrospect and prospect, NBER Working Paper 34621, National Bureau of Economic Research.

Campbell, John Y, Tarun Ramadorai, and Allie Schwartz, 2009, Caught on tape: Institutional trading, stock returns, and earnings announcements, *Journal of Financial Economics* 92, 66–91.

Chaudhary, Manav, Zhiyu Fu, and Jian Li, 2023, Corporate bond multipliers: Substitutes matter, *Available at SSRN* .

Chaudhry, Aditya, and Carter Davis, 2026, The origins of the factor zoo: Investors weakly substitute across stocks, *Available at SSRN 6545318* .

Cochrane, John H, 2008, The dog that did not bark: A defense of return predictability, *The Review of Financial Studies* 21, 1533–1575.

Conrad, Jennifer, Nishad Kapadia, and Yuhang Xing, 2014, Death and jackpot: Why do individual investors hold overpriced stocks?, *Journal of Financial Economics* 113, 455–475.

Cosemans, Mathijs, and Rik Frehen, 2021, Salience theory and stock prices: Empirical evidence, *Journal of Financial Economics* 140, 460–483.

Da, Zhi, Xing Huang, and Lawrence J. Jin, 2021, Extrapolative beliefs in the cross-section: What can we learn from the crowds?, *Journal of Financial Economics* 140, 175–196.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *Journal of Finance* 53, 1839–1885.

Daniel, Kent D., David Hirshleifer, and Avanidhar Subrahmanyam, 2001, Overconfidence, arbitrage, and equilibrium asset pricing, *Journal of Finance* 56, 921–965.

Davis, Carter, 2025, Classical versus demand system asset pricing, *Available at SSRN 5232345* .

Davis, Carter, Xiao Han, Stanislav Sokolinski, and Andrea Tamoni, 2025a, Rewriting expected returns, *Available at SSRN 5169378* .

Davis, Carter, Mahyar Kargar, and Jiacui Li, 2025b, Why do portfolio choice models predict inelastic demand?, *Journal of Financial Economics, Forthcoming* .

Davis, Carter, Mahyar Kargar, Jiacui Li, and Dejanir Silva, 2026, On the recovery of demand elasticities in dynamic settings, Working paper.

De Bondt, Werner FM, and Richard Thaler, 1985, Does the stock market overreact?, *The Journal of finance* 40, 793–805.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Noise Trader Risk in Financial Markets, *Journal of Political Economy* 98, 703–738.

Dew-Becker, Ian, Stefano Giglio, and Pooya Molavi, 2025, Learning and the emergence of nonlinearity in financial markets, NBER Working Paper 34584.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina, 2002, Differences of opinion and the cross section of stock returns, *Journal of Finance* 57, 2113–2141.

Dimson, Elroy, 1979, Risk measurement when shares are subject to infrequent trading, *Journal of Financial Economics* 7, 197–226.

Doeskeland, Trond M., and Hans K. Hvide, 2011, Do individual investors have asymmetric information based on work experience?, *Journal of Finance* 66, 1011–1041.

Fama, Eugene F, 1965, The behavior of stock-market prices, *The journal of Business* 38, 34–105.

Frazzini, Andrea, 2006, The disposition effect and underreaction to news, *The Journal of Finance* 61, 2017–2046.

Frazzini, Andrea, and Owen A Lamont, 2008, Dumb money: Mutual fund flows and the cross-section of stock returns, *Journal of Financial Economics* 88, 299–322.

Fuchs, William, Satoshi Fukuda, and Daniel Neuhann, 2023, Demand-system asset pricing: Theoretical foundations, *Available at SSRN 4672473* .

Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and H Eugene Stanley, 2006, Institutional Investors and Stock Market Volatility, *Quarterly Journal of Economics* 44.

Gabaix, Xavier, Ralph SJ Koijen, Federico Mainardi, Sangmin Oh, and Motohiro Yogo, 2022, Asset demand of U.S. households Working paper.

Gomes, Francisco, Michael Haliassos, and Tarun Ramadorai, 2021, Household finance, *Journal of Economic Literature* 59, 919–1000.

Greenwood, Robin, and Andrei Shleifer, 2014, Expectations of returns and expected returns, *The Review of Financial Studies* 27, 714–746.

Grinblatt, Mark, and Bing Han, 2005, Prospect theory, mental accounting, and momentum, *Journal of financial economics* 78, 311–339.

Grinblatt, Mark, Matti Keloharju, and Juhani T Linnainmaa, 2012, Iq, trading behavior, and performance, *Journal of Financial Economics* 104, 339–362.

Haddad, Valentin, Zhiguo He, Paul Huebner, Péter Kondor, and Erik Loualiche, 2025, Causal inference for asset pricing, *Available at SSRN 5187305* .

He, Zhiguo, Péter Kondor, and Jessica S Li, 2025, Demand elasticity in dynamic asset pricing, Working paper.

Hong, Harrison, and David A. Sraer, 2016, Speculative Betas, *The Journal of Finance* 71, 2095–2144.

Hong, Harrison, and Jeremy C Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *The Journal of finance* 54, 2143–2184.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Jensen, Theis Ingerslev, Bryan T. Kelly, and Lasse Heje Pedersen, 2023, Is there a replication crisis in finance?, *Journal of Finance* 78, 2465–2518.

Kaniel, Ron, Gideon Saar, and Sheridan Titman, 2008, Individual Investor Trading and Stock Returns, *The Journal of Finance* 63, 273–310, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2008.01316.x.

Kelley, Eric K., and Paul C. Tetlock, 2013, How Wise Are Crowds? Insights from Retail Orders and Stock Returns, *The Journal of Finance* 68, 1229–1265.

Kelly, Bryan T, Seth Pruitt, and Yinan Su, 2019, Characteristics are covariances: A unified model of risk and return, *Journal of Financial Economics* 134, 501–524.

Koijen, Ralph S J, Robert J Richmond, and Motohiro Yogo, 2022, Which Investors Matter for Equity Valuations and Expected Returns?, *Working Paper* 54.

Koijen, Ralph S. J., and Motohiro Yogo, 2019, A Demand System Approach to Asset Pricing, *Journal of Political Economy* .

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *Journal of Finance* 73, 1183–1223.

Kumar, Alok, 2009, Who gambles in the stock market?, *The Journal of Finance* 64, 1889–1933.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541–1578.

Li, Fulin, 2025, Retail trading and asset prices: The role of changing social dynamics, Working paper.

Li, Yizhang, Stanislav Sokolinski, and Andrea Tamoni, 2026, Which investors drive anomaly returns and how?, *Journal of Financial Economics* 179, 104257.

Lochstoer, Lars A., and Tyler Muir, 2026, Actively passive: The rise of market volatility, Working paper.

Loughran, Tim, and Jay R. Ritter, 1995, The new issues puzzle, *Journal of Finance* 50, 23–51.

Miller, Edward M., 1977, Risk, uncertainty, and divergence of opinion, *Journal of Finance* 32, 1151–1168.

Mitton, Todd, and Keith Vorkink, 2007, Equilibrium underdiversification and the preference for skewness, *Review of Financial Studies* 20, 1255–1288.

Odean, Terrance, 1998, Are investors reluctant to realize their losses?, *The Journal of Finance* 53, 1775–1798.

Odean, Terrance, 1999, Do investors trade too much?, *American Economic Review* 89, 1279–1298.

Pastor, Lubos, Robert Stambaugh, and Lucian A. Taylor, 2022, Dissecting green returns, *Journal of Financial Economics* 146, 403–424.

Polk, Christopher, and Paola Sapienza, 2009, The stock market and corporate investment: A test of catering theory, *Review of Financial Studies* 22, 187–217.

Scheinkman, José A., and Wei Xiong, 2003, Overconfidence and speculative bubbles, *Journal of Political Economy* 111, 1183–1219.

Shefrin, Hersh, and Meir Statman, 1985, The disposition to sell winners too early and ride losers too long: Theory and evidence, *The Journal of Finance* 40, 777–790.

Shleifer, Andrei, 2000, *Inefficient markets: An introduction to behavioural finance* (Oup Oxford).

Stambaugh, Robert F, Jianfeng Yu, and Yu Yuan, 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104, 288–302.

Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan, 2015, Arbitrage asymmetry and the idiosyncratic volatility puzzle, *Journal of Finance* 70, 1903–1948.

Stein, Jeremy C., 1996, Rational capital budgeting in an irrational world, *Journal of Finance* 51, 1111–1148.

van Binsbergen, Jules H., Benjamin David, and Christian C. Opp, 2025, How (not) to identify demand elasticities in dynamic asset markets, NBER Working Paper No. 34528.

van der Beck, Philippe, 2022, Short-versus long-run demand elasticities in asset pricing, *Asset Pricing (May 16, 2022). Swiss Finance Institute Research Paper* .

van der Beck, Philippe, Lorenzo Bretscher, and Julie Zhiyu Fu, 2025, A bound on price impact and disagreement, Swiss Finance Institute Research Paper No. 25-77.

Van der Beck, Philippe, Coralie Jaunin, and Cameron Cohen, 2026, The equity market implications of the retail investment boom, *Management Science* Fortcoming.

Weber, Michael, 2018, Cash flow duration and the term structure of equity returns, *Journal of Financial Economics* 130, 325–344.

Zhang, X. Frank, 2006, Information uncertainty and stock returns, *The Journal of Finance* 61, 105–137.

A Anomaly Portfolio Construction and Coverage

We begin with the 153 stock characteristics in the Global Factor Data library of Jensen et al. (2023) that can be linked to the Norwegian inside-stock panel. We retain a characteristic if it has at least 85 months with a valid long–short return, at least 98 months with distinct 30th and 70th percentile breakpoints, and nonempty long and short legs in every retained month. These rules retain 139 characteristics. Return support ranges from 85 to 119 months, with a median of 119.

For every retained characteristic, we use the published sign convention in the Global Factor Data documentation. Each month we form equal-weighted long and short legs using the 30th and 70th percentiles among inside stocks with valid lagged signals. This construction holds the stock universe, portfolio rule, and weighting scheme fixed across all 556 behavioral-tilt–anomaly cells. Table 12 gives the exact code, definition, long-side orientation, anomaly family, and original citation for every retained characteristic. Table 13 reports the 14 excluded characteristics and their realized support.

Table 12. Definitions of the 139 anomaly portfolios

Code	JKP definition	Long side	Family	Original citation
aliq_at	Liquidity of book assets	Low	Intangibles	Ortiz-Molina and Phillips (2014)
aliq_mat	Liquidity of market assets	Low	Intangibles	Ortiz-Molina and Phillips (2014)
ami_126d	Amihud Measure	High	Trading Frictions	Amihud (2002)
at_be	Book leverage	Low	Profitability	Fama and French (1992)
at_gr1	Asset Growth	Low	Investment	Cooper Gulen and Schill (2008)
at_me	Assets-to-market	High	Value	Fama and French (1992)
at_turnover	Capital turnover	High	Profitability	Haugen and Baker (1996)
be_gr1a	Change in common equity	Low	Investment	Richardson et al. (2005)
be_me	Book-to-market equity	High	Value	Rosenberg Reid and Lanstein (1985)
beta_60m	Market Beta	Low	Trading Frictions	Fama and MacBeth (1973)
beta_dimson_21d	Dimson beta	Low	Trading Frictions	Dimson (1979)
betabab_1260d	Frazzini-Pedersen market beta	Low	Trading Frictions	Frazzini and Pedersen (2014)
betadown_252d	Downside beta	Low	Trading Frictions	Ang Chen and Xing (2006)
bev_mev	Book-to-market enterprise value	High	Value	Penman Richardson and Tuna (2007)
bidaskhl_21d	The high-low bid-ask spread	High	Trading Frictions	Corwin and Schultz (2012)
capex_abn	Abnormal corporate investment	Low	Investment	Titman Wei and Xie (2004)
capx_gr1	CAPEX growth (1 year)	Low	Investment	Xie (2001)
capx_gr2	CAPEX growth (2 years)	Low	Investment	Anderson and Garcia-Feijoo (2006)
capx_gr3	CAPEX growth (3 years)	Low	Investment	Anderson and Garcia-Feijoo (2006)
cash_at	Cash-to-assets	High	Intangibles	Palazzo (2012)
coa_gr1a	Change in current operating assets	Low	Investment	Richardson et al. (2005)

Code	JKP definition	Long side	Family	Original citation
col_gr1a	Change in current operating liabilities	Low	Investment	Richardson et al. (2005)
cop_at	Cash-based operating profits-to-book assets	High	Profitability	
cop_atl1	Cash-based operating profits-to-lagged book assets	High	Profitability	Ball et al. (2016)
corr_1260d	Market correlation	Low	New	Assness, Frazzini, Gormsen, Pedersen (2020)
coskew_21d	Coskewness	Low	Trading Frictions	Harvey and Siddique (2000)
cowc_gr1a	Change in current operating working capital	Low	Investment	Richardson et al. (2005)
dbnetis_at	Net debt issuance	Low	Investment	Bradshaw Richardson and Sloan (2006)
debt_gr3	Growth in book debt (3 years)	Low	Investment	Lyandres Sun and Zhang (2008)
debt_me	Debt-to-market	High	Value	Bhandari (1988)
dgp_dsale	Change gross margin minus change sales	High	Intangibles	Abarbanell and Bushee (1998)
div12m_me	Dividend yield	High	Value	Litzenberger and Ramaswamy (1979)
dolvol_126d	Dollar trading volume	Low	Trading Frictions	Brennan Chordia and Subrahmanyam (1998)
dolvol_var_126d	Coefficient of variation for dollar trading volume	Low	Trading Frictions	Chordia Subrahmanyam and Anshuman (2001)
dsale_dinv	Change sales minus change Inventory	High	Intangibles	Abarbanell and Bushee (1998)
dsale_drec	Change sales minus change receivables	Low	Intangibles	Abarbanell and Bushee (1998)
dsale_dsga	Change sales minus change SG&A	High	Intangibles	Abarbanell and Bushee (1998)
earnings_variability	Earnings variability	Low	Intangibles	Francis et al. (2004)
ebit_bev	Return on net operating assets	High	Profitability	Soliman (2008)
ebit_sale	Profit margin	High	Profitability	Soliman (2008)

Code	JKP definition	Long side	Family	Original citation
ebitda_mev	Ebitda-to-market enterprise value	High	Value	Loughran and Wellman (2011)
emp_gr1	Hiring rate	Low	Intangibles	Belo Lin and Bazdresch (2014)
eq_dur	Equity duration	Low	Value	Dechow Sloan and Soliman (2004)
eqnetis_at	Net equity issuance	Low	Investment	Bradshaw Richardson and Sloan (2006)
eqnpo_12m	Equity net payout	High	Investment	Daniel and Titman (2006)
eqnpo_me	Net payout yield	High	Value	Boudoukh et al. (2007)
f_score	Pitroski F-score	High	Profitability	Piotroski (2000)
fcf_me	Free cash flow-to-price	High	Value	Lakonishok Shleifer and Vishny (1994)
fnl_gr1a	Change in financial liabilities	Low	Investment	Richardson et al. (2005)
gp_at	Gross profits-to-assets	High	Profitability	Novy-Marx (2013)
gp_atl1	Gross profits-to-lagged assets	High	Profitability	
inv_gr1	Inventory growth	Low	Investment	Belo and Lin (2011)
inv_gr1a	Inventory change	Low	Investment	Thomas and Zhang (2002)
iskew_capm_21d	Idiosyncratic skewness from the CAPM	Low	Trading Frictions	
ival_me	Intrinsic value-to-market	High	Value	Frankel and Lee (1998)
ivol_capm_21d	Idiosyncratic volatility from the CAPM (21 days)	Low	Trading Frictions	
ivol_capm_252d	Idiosyncratic volatility from the CAPM (252 days)	Low	Trading Frictions	Ali Hwang and Trombley (2003)
kz_index	Kaplan-Zingales index	High	Intangibles	Lamont Polk and Saa-Requejo (2001)
lnoa_gr1a	Change in long-term net operating assets	Low	Investment	Fairfield Whisenant and Yohn (2003)
lti_gr1a	Change in long-term investments	Low	Investment	Richardson et al. (2005)
market_equity	Market Equity	Low	Trading Frictions	Banz (1981)
mispricing_mgmt	Mispricing factor: Management	High	New	Stambaugh and Yuan (2016)

Code	JKP definition	Long side	Family	Original citation
mispricing_perf	Mispricing factor: Performance	High	New	Stambaugh and Yuan (2016)
ncoa_gr1a	Change in noncurrent operating assets	Low	Investment	Richardson et al. (2005)
ncol_gr1a	Change in noncurrent operating liabilities	Low	Investment	Richardson et al. (2005)
netdebt_me	Net debt-to-price	Low	Value	Penman Richardson and Tuna (2007)
netis_at	Net total issuance	Low	Investment	Bradshaw Richardson and Sloan (2006)
nfna_gr1a	Change in net financial assets	High	Investment	Richardson et al. (2005)
ni_ar1	Earnings persistence	High	Intangibles	Francis et al. (2004)
ni_be	Return on equity	High	New	Haugen and Baker (1996)
ni_inc8q	Number of consecutive quarters with earnings increases	High	Momentum	Barth Elliott and Finn (1999)
ni_ivol	Earnings volatility	High	Intangibles	Francis et al. (2004)
ni_me	Earnings-to-price	High	Value	Basu (1983)
niq_at	Quarterly return on assets	High	Profitability	Balakrishnan Bartov and Faurel (2010)
niq_at_chg1	Change in quarterly return on assets	High	Profitability	
niq_be	Quarterly return on equity	High	Profitability	Hou Xue and Zhang (2015)
niq_be_chg1	Change in quarterly return on equity	High	Profitability	
niq_su	Standardized earnings surprise	High	Momentum	Foster Olsen and Shevlin (1984)
nncoa_gr1a	Change in net noncurrent operating assets	Low	Investment	Richardson et al. (2005)
noa_at	Net operating assets	Low	Investment	Hirshleifer et al. (2004)
noa_gr1a	Change in net operating assets	Low	Investment	Hirshleifer et al. (2004)
o_score	Ohlson O-score	Low	Profitability	Dichev (1998)
oaccruals_at	Operating accruals	Low	Investment	Sloan (1996)
oaccruals_ni	Percent operating accruals	Low	Investment	Hafzalla Lundholm and Van Winkle (2011)

Code	JKP definition	Long side	Family	Original citation
ocf_at	Operating cash flow to assets	High	New	Bouchard, Krüger, Landier and Thesmar (2019)
ocf_at_chg1	Change in operating cash flow to assets	High	New	Bouchard, Krüger, Landier and Thesmar (2019)
ocf_me	Operating cash flow-to-market	High	Value	Desai Rajgopal and Venkatachalam (2004)
ocfq_saleq_std	Cash flow volatility	Low	Intangibles	Huang (2009)
op_at	Operating profits-to-book assets	High	Profitability	
op_atl1	Operating profits-to-lagged book assets	High	Profitability	Ball et al. (2016)
ope_be	Operating profits-to-book equity	High	Profitability	Fama and French (2015)
ope_bel1	Operating profits-to-lagged book equity	High	Profitability	
opex_at	Operating leverage	High	Intangibles	Novy-Marx (2011)
pi_nix	Taxable income-to-book income	High	Profitability	Lev and Nissim (2004)
ppeinv_gr1a	Change PPE and Inventory	Low	Investment	Lyandres Sun and Zhang (2008)
prc	Price per share	Low	Trading Frictions	Miller and Scholes (1982)
prc_highprc_252d	Current price to high price over last year	High	Momentum	George and Hwang (2004)
qmj	Quality minus Junk: Composite	High	New	Assness, Frazzini and Pedersen (2018)
qmj_growth	Quality minus Junk: Growth	High	New	Assness, Frazzini and Pedersen (2018)
qmj_prof	Quality minus Junk: Profitability	High	New	Assness, Frazzini and Pedersen (2018)
qmj_safety	Quality minus Junk: Safety	High	New	Assness, Frazzini and Pedersen (2018)
rd_me	R&D-to-market	High	Intangibles	Chan Lakonishok and Sougiannis (2001)

Code	JKP definition	Long side	Family	Original citation
rd_sale	R&D-to-sales	High	Intangibles	Chan Lakonishok and Sougiannis (2001)
ret_12_1	Price momentum t-12 to t-1	High	Momentum	Jegadeesh and Titman (1993)
ret_12_7	Price momentum t-12 to t-7	High	New	Novy-Marx (2012)
ret_1_0	Short-term reversal	Low	Trading Frictions	Jegadeesh (1990)
ret_3_1	Price momentum t-3 to t-1	High	New	Jegedeesh and Titman (1993)
ret_60_12	Long-term reversal	Low	Value	De Bondt and Thaler (1985)
ret_6_1	Price momentum t-6 to t-1	High	Momentum	Jegadeesh and Titman (1993)
ret_9_1	Price momentum t-9 to t-1	High	New	Jegedeesh and Titman (1993)
rmax1_21d	Maximum daily return	Low	Trading Frictions	Bali Cakici and Whitelaw (2011)
rmax5_21d	Highest 5 days of return	Low	New	Bali, Brown, Murray and Tang (2017)
rmax5_rvol_21d	Highest 5 days of return scaled by volatility	Low	New	Assness, Frazzini, Gormsen, Pedersen (2020)
rskew_21d	Total skewness	Low	Trading Frictions	Bali Engle and Murray (2016)
rvol_21d	Return volatility	Low	Trading Frictions	Ang et al. (2006)
sale_bev	Asset turnover	High	Profitability	Soliman (2008)
sale_emp_gr1	Labor force efficiency	High	Intangibles	Abarbanell and Bushee (1998)
sale_gr1	Sales Growth (1 year)	Low	Value	Lakonishok Shleifer and Vishny (1994)
sale_gr3	Sales Growth (3 years)	Low	Value	Lakonishok Shleifer and Vishny (1994)
sale_me	Sales-to-market	High	Value	Barbee Mukherji and Raines (1996)
saleq_gr1	Sales growth (1 quarter)	Low	Profitability	
saleq_su	Standardized Revenue surprise	High	Momentum	Jegadeesh and Livnat (2006)
seas_1_1an	Year 1-lagged return, annual	High	Intangibles	Heston and Sadka (2008)
seas_1_1na	Year 1-lagged return, nonannual	High	Intangibles	Heston and Sadka (2008)
seas_2_5an	Years 2-5 lagged returns, annual	High	Intangibles	Heston and Sadka (2008)

Code	JKP definition	Long side	Family	Original citation
seas_2_5na	Years 2-5 lagged returns, nonannual	Low	Intangibles	Heston and Sadka (2008)
seas_6_10an	Years 6-10 lagged returns, annual	High	Intangibles	Heston and Sadka (2008)
seas_6_10na	Years 6-10 lagged returns, nonannual	Low	Intangibles	Heston and Sadka (2008)
sti_gr1a	Change in short-term investments	High	Investment	Richardson et al. (2005)
taccruals_at	Total accruals	Low	Investment	Richardson et al. (2005)
taccruals_ni	Percent total accruals	Low	Investment	Hafzalla Lundholm and Van Winkle (2011)
tangibility	Asset tangibility	High	Intangibles	Hahn and Lee (2009)
tax_gr1a	Tax expense surprise	High	Momentum	Thomas and Zhang (2011)
turnover_126d	Share turnover	Low	Trading Frictions	Datar Naik and Radcliffe (1998)
turnover_var_126d	Coefficient of variation for share turnover	Low	Trading Frictions	Chordia Subrahmanyam and Anshuman (2001)
z_score	Altman Z-score	High	Profitability	Dichev (1998)
zero_trades_126d	Number of zero trades with turnover as tiebreaker (6 months)	High	Trading Frictions	Liu (2006)
zero_trades_21d	Number of zero trades with turnover as tiebreaker (1 month)	High	Trading Frictions	Liu (2006)
zero_trades_252d	Number of zero trades with turnover as tiebreaker (12 months)	High	Trading Frictions	Liu (2006)

Table 13. Characteristics excluded by the coverage rule

Code	Name	Return support	Distinct sorts	Reason
age	Firm age	0	0	insufficient distinct 30/70 breakpoints
chcsho_12m	Net stock issues	42	42	insufficient distinct 30/70 breakpoints
eqpo_me	Payout yield	63	63	fewer than minimum raw-support months
iskew_ff3_21d	Idiosyncratic skewness from the Fama-French 3-factor model	46	46	fewer than minimum raw-support months
iskew_hxz4_21d	Idiosyncratic skewness from the q-factor model	54	54	fewer than minimum raw-support months
ivol_ff3_21d	Idiosyncratic volatility from the Fama-French 3-factor model	46	46	fewer than minimum raw-support months
ivol_hxz4_21d	Idiosyncratic volatility from the q-factor model	54	54	fewer than minimum raw-support months
rd5_at	R&D capital-to-book assets	11	11	fewer than minimum raw-support months
resff3_12_1	Residual momentum t-12 to t-1	35	35	fewer than minimum raw-support months
resff3_6_1	Residual momentum t-6 to t-1	33	33	fewer than minimum raw-support months
seas_11_15an	Years 11-15 lagged returns, annual	75	75	fewer than minimum raw-support months
seas_11_15na	Years 11-15 lagged returns, nonannual	66	66	fewer than minimum raw-support months
seas_16_20an	Years 16-20 lagged returns, annual	3	3	fewer than minimum raw-support months
seas_16_20na	Years 16-20 lagged returns, nonannual	0	0	fewer than minimum raw-support months

B Theoretical predictions for biases and anomalies

In this appendix, we explain the anomaly predictions we test in the paper: which biases generate which anomalies. The predictions are derived from the existing literature. We start from the literature surveyed in Barberis (2018). For each bias, or empirical measure of a bias, we list the economic mechanisms through which the bias can generate anomaly returns. Because we use a comprehensive universe of 139 anomalies, we then explain how we map each mechanism to the anomalies in our counterfactual exercise. This mapping produces 91 predictions spanning 51 anomalies. Table 14 lists the economic mechanisms. Within each bias, the mechanisms run from the earliest paper to the most recent. The following sections discuss each mechanism in turn, and Section B.5 records the anomaly-level assignment.

Table 14. Index of predictions

Bias	Code	Anomaly category	# of Anomalies	Economic argument
Overconfidence	O1	Momentum	6	Biased self-attribution generates momentum.
	O2	Post-news drift	3	Biased self-attribution generates post-earnings drift.
	O3	Investment	10	Firms invest and issue equity when overpriced.
	O4	Value	10	Overconfidence overprices low fundamental-to-price stocks.
	O5	Duration/Growth	3	Overconfidence overprices high sales-growth and long-duration stocks.
	O6	Volatility	5	Overconfidence inflates prices and raises price volatility.
	O7	Turnover	2	Overconfidence inflates prices and raises trading volume.
	O8	Market beta	3	Disagreement under short-sale constraints overprices high-beta stocks.
Lottery preference	L1	Skewness	3	Probability weighting overprices positively skewed stocks.
	L2	Issuance	3	Prospect theory overprices equity issuers because their returns are more positively skewed.
	L3	Maximum returns	3	Probability weighting overprices stocks with extreme daily gains.
	L4	Volatility	5	Controlling for maximum returns explains the low returns on volatile stocks.
	L5	Value	10	Salience overprices growth stocks whose large upside stands out.
	L6	Distress and safety	3	Distressed stocks have lottery-like payoffs and are overpriced.
	L7	Market beta	3	Lottery demand overprices high-beta stocks.
Disposition effect	D1	Momentum	6	Selling winners and holding losers generates momentum.
	D2	Post-news drift	3	Selling winners and holding losers slows the incorporation of news.
Extrapolation	E1	Value	10	Extrapolation of past growth overprices glamour stocks.

B.1 Overconfidence

B.1.1 O1: Momentum

Daniel et al. (1998) link overconfidence with biased self-attribution to momentum. Public news that confirms a private signal raises confidence by more than disconfirming news lowers it, so the price continues in the direction of the initial overreaction. The result is short-run momentum.[15] We apply this prediction to momentum anomalies.

Anomalies covered: `prc_highprc_252d`, `ret_12_1`, `ret_12_7`, `ret_3_1`, `ret_6_1`, and `ret_9_1`.

B.1.2 O2: Post-news drift

In the model of Daniel et al. (1998), a public signal that confirms the private information triggers further overreaction before the price corrects. They argue that short-run post-earnings drift is this continuation. We apply this prediction to post-news momentum.

Anomalies covered: `niq_su`, `saleq_su`, and `tax_gr1a`.

B.1.3 O3: Corporate investment and issuance

Daniel et al. (1998) link investor overconfidence to abnormal returns following corporate events such as issuance. "If investors are overconfident, then selective events that are initiated when the stock is undervalued (overvalued) by the market will on average be associated with positive (negative) announcement-date abnormal price changes and will on average be followed by positive (negative) post-announcement abnormal price changes." "We expect firms to issue securities when they believe their stocks are overvalued."[16] Stein (1996) shows that a similar mechanism raises investment. "If the manager is interested in maximizing the current stock price, he must cater to any misperceptions that investors have. Thus if investors are overly optimistic about the prospects for the firm's assets ... the manager should be willing to invest very aggressively in these assets." Baker, Stein, and Wurgler (2003) find that the investment of equity-dependent firms responds to stock prices, and Polk

[15] Jegadeesh and Titman (1993) document the return premium.

[16] Loughran and Ritter (1995) document these low returns. Baker and Wurgler (2002) find that firms issue equity when valuations are high.

and Sapienza (2009) find that firms with high abnormal investment earn low subsequent returns.

We apply this argument to investment and to equity issuance.

Anomalies covered: `at_gr1`, `be_gr1a`, `capex_abn`, `capx_gr1`, `capx_gr2`, `capx_gr3`, `emp_gr1`, `eqnetis_at`, `eqnpo_12m`, and `netis_at`.

B.1.4 O4: Value

Daniel, Hirshleifer, and Subrahmanyam (2001) study investors who are overconfident about private information: they treat a private signal as more precise than it is.[17] Prices then overreact. After a favorable signal the price is too high, and after an unfavorable signal it is too low. A fundamental-to-price ratio places price in the denominator, so the ratio is low when the stock is overpriced and high when it is underpriced. High ratios therefore predict high subsequent returns.

Daniel et al. (2001) argue that this mispricing is captured by fundamental-to-price ratios, so anomalies that sort on those ratios are subject to the same argument. We use the ten fundamental-to-price characteristics in the value category.

Anomalies covered: `at_me`, `be_me`, `bev_mev`, `div12m_me`, `ebitda_mev`, `fcf_me`, `ival_me`, `ni_me`, `ocf_me`, and `sale_me`.

B.1.5 O5: Duration and growth

Daniel et al. (2001) argue that overconfidence overprices growth stocks. Similarly, Daniel et al. (1998) argue that the overconfidence-driven mispricing is stronger for growth stocks whose value depends on hard-to-value growth options. Past sales growth and equity duration identify that growth side, so high sales growth and long duration are the overpriced stocks. Weber (2018) refers to the overconfidence mechanism of Daniel et al. (1998): overconfidence creates stronger mispricing when valuation requires judgment and feedback is slow. Long-duration stocks are this case, because their cash flows arrive far in the future. Analysts

[17] Barber and Odean (2001) show that men, whom they associate with greater overconfidence, trade more, and Betermier, Calvet, and Sodini (2017) show that men tilt toward growth rather than value.

extrapolate their high past growth. The low returns are larger after periods of high investor sentiment and are concentrated among short-sale-constrained stocks.

Anomalies covered: `sale_gr1`, `sale_gr3`, and `eq_dur`.

B.1.6 O6: Volatility

Scheinkman and Xiong (2003) link overconfidence to asset bubbles and volatility. "Overconfidence generates disagreements among agents regarding asset fundamentals. With short-sale constraints, an asset buyer acquires an option to sell the asset to other agents when those agents have more optimistic beliefs. ... This causes a significant bubble component in asset prices. ... In equilibrium, bubbles are accompanied by large trading volume and high price volatility."[18] Relatedly, Zhang (2006) refers to the overconfidence mechanism of Daniel et al. (1998, 2001) and argues that biases such as overconfidence are stronger when stocks are hard to value. He uses return volatility and cash-flow volatility as proxies for hard-to-value stocks.

We apply these predictions to anomalies based on return and fundamental volatility.

Anomalies covered: `earnings_variability`, `ivol_capm_21d`, `ivol_capm_252d`, `ocfq_saleq_std`, and `rvol_21d`.

B.1.7 O7: Turnover

In the model of Scheinkman and Xiong (2003), overconfidence also produces high trading volume. We apply this prediction to turnover anomalies.

Anomalies covered: `turnover_126d` and `turnover_var_126d`.

B.1.8 O8: Market beta

Hong and Sraer (2016) argue that high-beta stocks are overpriced when investors disagree about the market. "When investors disagree about the stock market's prospects, high-beta assets are more sensitive to this aggregate disagreement, experience greater divergence of opinion about their payoffs, and are overpriced due to short-sales constraints." "Macro-

[18] Diether et al. (2002) find that stocks about which analysts disagree more earn low subsequent returns.

disagreement might emanate from many sources such as heterogeneous priors or cognitive biases such as overconfidence." We apply this mechanism to the beta anomalies.

Anomalies covered: `beta_60m`, `beta_dimson_21d`, and `betabab_1260d`.

B.2 Lottery preference

B.2.1 L1: Skewness

Barberis and Huang (2008) show that cumulative prospect theory, through its probability weighting component, gives investors a preference for lottery-like payoffs. A positively skewed security can then be overpriced and can earn a negative average excess return. Boyer et al. (2010) find that stocks with high expected idiosyncratic skewness earn low subsequent returns.

Anomalies covered: `coskew_21d`, `iskew_capm_21d`, and `rskew_21d`.

B.2.2 L2: Equity issuance

Barberis et al. (2021) find that prospect theory helps explain the low returns of equity issuers. In their sample, the low-return leg of issuance anomalies is more positively skewed, and probability weighting leads investors to accept a lower average return on that leg. Barberis and Huang (2008) make the same argument for initial public offerings, whose value is largely a growth option and whose returns are positively skewed.

Anomalies covered: `eqnetis_at`, `eqnpo_12m`, and `netis_at`.

B.2.3 L3: Maximum returns

Bali et al. (2011) use the maximum daily return as the measure of a stock's lottery payoff. Stocks with a high maximum daily return over the past month earn low subsequent returns, and this maximum return accounts for the idiosyncratic-volatility effect. Barberis et al. (2021) find that prospect theory also helps explain the maximum-daily-return anomaly because the low-return leg is more positively skewed. We apply this prediction to maximum-return anomalies.

Anomalies covered: `rmax1_21d`, `rmax5_21d`, and `rmax5_rvol_21d`.

B.2.4 L4: Volatility

The same mechanism of Bali et al. (2011) links lottery preferences to volatility anomalies, as controlling for maximum returns explains the low returns on volatile stocks. Barberis et al. (2021) find that prospect theory helps explain the idiosyncratic-volatility anomaly because the low-return leg is more positively skewed, and this effect leads investors to accept a lower average return on it.

Anomalies covered: `earnings_variability`, `ivol_capm_21d`, `ivol_capm_252d`, `ocfq_saleq_std`, and `rvol_21d`.

B.2.5 L5: Value

Bordalo et al. (2013) argue that investors prefer assets with a salient chance of a very high payoff, and they apply that preference to the growth-value puzzle. Growth stocks are overpriced because their large upside is salient. Value stocks are underpriced because a large downside, such as bankruptcy, is salient. We apply the salience argument to the fundamental-to-price characteristics.

Anomalies covered: `at_me`, `be_me`, `bev_mev`, `div12m_me`, `ebitda_mev`, `fcf_me`, `ival_me`, `ni_me`, `ocf_me`, and `sale_me`.

B.2.6 L6: Distress and safety

Conrad, Kapadia, and Xing (2014) show that firms with a high probability of default also tend to have a high probability of a jackpot payoff, and that stocks with a high jackpot probability earn low returns. Stocks with high default risk but a low jackpot probability do not. Barberis et al. (2021) find that prospect theory prices failure-probability and O-score portfolios because the low-return leg is more positively skewed. We apply the argument to distress and safety.

Anomalies covered: `z_score`, `o_score`, and `qmj_safety`.

B.2.7 L7: Market beta

Bali et al. (2017) argue that lottery demand accounts for the beta anomaly. The low returns of high-beta stocks are concentrated where lottery demand has its price impact, and the anomaly is absent once portfolios are neutralized to lottery demand.

Anomalies covered: `beta_60m`, `beta_dimson_21d`, and `betabab_1260d`.

B.3 Disposition effect

B.3.1 D1: Momentum

Grinblatt and Han (2005) argue that the disposition effect can generate momentum.[19] "The tendency of some investors to hold on to their losing stocks, driven by prospect theory and mental accounting, creates a spread between a stock's fundamental value and its equilibrium price, as well as price underreaction to information. Spread convergence, arising from the random evolution of fundamental values and updating of reference prices, generates predictable equilibrium prices that will be interpreted as possessing momentum." Barberis et al. (2021) find that prospect theory helps explain the momentum anomaly. We apply this prediction to momentum anomalies.

Anomalies covered: `prc_highprc_252d`, `ret_12_1`, `ret_12_7`, `ret_3_1`, `ret_6_1`, and `ret_9_1`.

B.3.2 D2: Post-news drift

Frazzini (2006) links the same mechanism to news momentum. "This paper tests whether the 'disposition effect,' that is the tendency of investors to ride losses and realize gains, induces 'underreaction' to news, leading to return predictability." Barberis et al. (2021) find that prospect theory helps explain post-earnings announcement drift. We apply this prediction to news momentum anomalies.

Anomalies covered: `niq_su`, `saleq_su`, and `tax_gr1a`.

[19] The trading pattern is documented by Shefrin and Statman (1985) and Odean (1998).

B.4 Extrapolation

B.4.1 E1: Value

Lakonishok et al. (1994) link extrapolation of past performance to the value anomaly.[20][21] "Value strategies might produce higher returns because they are contrarian to 'naive' strategies followed by other investors. These naive strategies might range from extrapolating past earnings growth too far into the future, to assuming a trend in stock prices, to overreacting to good or bad news. ...Some investors tend to get overly excited about stocks that have done very well in the past and buy them up, so that these 'glamour' stocks become overpriced. Similarly, they overreact to stocks that have done very badly, oversell them, and these out-of-favor 'value' stocks become underpriced." "Our version of the contrarian model predicts that differences in expected future growth rates are linked to past growth and overestimate actual future growth differences between glamour and value firms." We apply the argument to the fundamental-to-price characteristics.

Anomalies covered: `at_me`, `be_me`, `bev_mev`, `div12m_me`, `ebitda_mev`, `fcf_me`, `ival_me`, `ni_me`, `ocf_me`, and `sale_me`.

B.5 Feature-level prediction index

The table below lists the 91 creator predictions across 51 anomaly characteristics. Every other cell has no directional prediction. The feature code is the exact identifier used to construct the anomaly portfolio.

[20] Our empirical bias measure captures extrapolation of returns, not extrapolation of fundamentals. Related work links extrapolation of fundamentals to anomalies (Barberis, Shleifer, and Vishny, 1998b; Lakonishok et al., 1994; Bordalo, Gennaioli, La Porta, and Shleifer, 2019). See also Barberis and Shleifer (2003) on the extrapolation of styles. We leave the exploration of the role of fundamental extrapolation to future research.

[21] The literature has also explored the link between extrapolation of returns and momentum. In the survey of Barberis (2018), extrapolation can generate momentum when beliefs exclude the most recent price change, whereas in Da, Huang, and Jin (2021) extrapolation leads to lower returns: investors extrapolate recent stock returns, and those beliefs negatively predict returns over the next week.

Anomaly characteristic	Overconfidence	Lottery	Disposition	Extrapolation
Abnormal corporate investment (capex_abn)	03	–	–	–
Altman Z-score (z_score)	–	L6	–	–
Asset Growth (at_gr1)	03	–	–	–
Assets-to-market (at_me)	04	L5	–	E1
Book-to-market enterprise value (bev_mev)	04	L5	–	E1
Book-to-market equity (be_me)	04	L5	–	E1
CAPEX growth (1 year) (capx_gr1)	03	–	–	–
CAPEX growth (2 years) (capx_gr2)	03	–	–	–
CAPEX growth (3 years) (capx_gr3)	03	–	–	–
Cash flow volatility (ocfq_saleq_std)	06	L4	–	–
Change in common equity (be_gr1a)	03	–	–	–
Coefficient of variation for share turnover (turnover_var_126d)	07	–	–	–
Coskewness (coskew_21d)	–	L1	–	–
Current price to high price over last year (prc_highprc_252d)	01	–	D1	–
Dimson beta (beta_dimson_21d)	08	L7	–	–
Dividend yield (div12m_me)	04	L5	–	E1
Earnings variability (earnings_variability)	06	L4	–	–
Earnings-to-price (ni_me)	04	L5	–	E1
Ebitda-to-market enterprise value (ebitda_mev)	04	L5	–	E1
Equity duration (eq_dur)	05	–	–	–
Equity net payout (eqnpo_12m)	03	L2	–	–
Frazzini-Pedersen market beta (betabab_1260d)	08	L7	–	–
Free cash flow-to-price (fcf_me)	04	L5	–	E1

Continued on next page

Anomaly characteristic	Overconfidence	Lottery	Disposition	Extrapolation
Highest 5 days of return (rmax5_21d)	–	L3	–	–
Highest 5 days of return scaled by volatility (rmax5_rvol_21d)	–	L3	–	–
Hiring rate (emp_gr1)	O3	–	–	–
Idiosyncratic skewness from the CAPM (iskew_capm_21d)	–	L1	–	–
Idiosyncratic volatility from the CAPM (21 days) (ivol_capm_21d)	O6	L4	–	–
Idiosyncratic volatility from the CAPM (252 days) (ivol_capm_252d)	O6	L4	–	–
Intrinsic value-to-market (ival_me)	O4	L5	–	E1
Market Beta (beta_60m)	O8	L7	–	–
Maximum daily return (rmax1_21d)	–	L3	–	–
Net equity issuance (eqnetis_at)	O3	L2	–	–
Net total issuance (netis_at)	O3	L2	–	–
Ohlson O-score (o_score)	–	L6	–	–
Operating cash flow-to-market (ocf_me)	O4	L5	–	E1
Price momentum t-12 to t-1 (ret_12_1)	O1	–	D1	–
Price momentum t-12 to t-7 (ret_12_7)	O1	–	D1	–
Price momentum t-3 to t-1 (ret_3_1)	O1	–	D1	–
Price momentum t-6 to t-1 (ret_6_1)	O1	–	D1	–
Price momentum t-9 to t-1 (ret_9_1)	O1	–	D1	–
Quality minus Junk: Safety (qmj_safety)	–	L6	–	–
Return volatility (rvol_21d)	O6	L4	–	–
Sales Growth (1 year) (sale_gr1)	O5	–	–	–
Sales Growth (3 years) (sale_gr3)	O5	–	–	–
Sales-to-market (sale_me)	O4	L5	–	E1
Share turnover (turnover_126d)	O7	–	–	–
Standardized earnings surprise (niq_su)	O2	–	D2	–

Continued on next page

Anomaly characteristic	Overconfidence	Lottery	Disposition	Extrapolation
Standardized Revenue surprise (saleq_su)	O2	–	D2	–
Tax expense surprise (tax_gr1a)	O2	–	D2	–
Total skewness (rskew_21d)	–	L1	–	–

C Anomaly-Category Definitions and Estimates

Figure 6 contains 18 tilt–category averages formed from 12 economic categories and 51 unique anomalies. Each bar gives every anomaly in its category equal weight. Category overlap is explicit: issuance is a three-anomaly subset of investment, and the same category can be shown for more than one behavioral tilt. Table 16 gives the complete membership.

Table 16. Anomaly-category definitions

Category	Anomaly characteristics
Market beta	Market Beta (beta_60m); Dimson beta (beta_dimson_21d); Frazzini-Pedersen market beta (betabab_1260d)
Duration/Growth	Sales Growth (1 year) (sale_gr1); Sales Growth (3 years) (sale_gr3); Equity duration (eq_dur)
Value	Assets-to-market (at_me); Book-to-market equity (be_me); Book-to-market enterprise value (bev_mev); Dividend yield (div12m_me); Ebitda-to-market enterprise value (ebitda_mev); Free cash flow-to-price (fcf_me); Intrinsic value-to-market (ival_me); Earnings-to-price (ni_me); Operating cash flow-to-market (ocf_me); Sales-to-market (sale_me)
Volatility	Idiosyncratic volatility from the CAPM (21 days) (ivol_capm_21d); Idiosyncratic volatility from the CAPM (252 days) (ivol_capm_252d); Return volatility (rvol_21d); Earnings variability (earnings_variability); Cash flow volatility (ocfq_saleq_std)

Category	Anomaly characteristics
Investment	Asset Growth (at_gr1); Change in common equity (be_gr1a); Abnormal corporate investment (capex_abn); CAPEX growth (1 year) (capx_gr1); CAPEX growth (2 years) (capx_gr2); CAPEX growth (3 years) (capx_gr3); Hiring rate (emp_gr1); Net equity issuance (eqnetis_at); Equity net payout (eqnpo_12m); Net total issuance (netis_at)
Momentum	Price momentum t-3 to t-1 (ret_3_1); Price momentum t-6 to t-1 (ret_6_1); Price momentum t-9 to t-1 (ret_9_1); Price momentum t-12 to t-1 (ret_12_1); Price momentum t-12 to t-7 (ret_12_7); Current price to high price over last year (prc_highprc_252d)
Post-news drift	Standardized earnings surprise (niq_su); Standardized Revenue surprise (saleq_su); Tax expense surprise (tax_gr1a)
Turnover	Share turnover (turnover_126d); Coefficient of variation for share turnover (turnover_var_126d)
Maximum returns	Maximum daily return (rmax1_21d); Highest 5 days of return (rmax5_21d); Highest 5 days of return scaled by volatility (rmax5_rvol_21d)
Issuance	Net equity issuance (eqnetis_at); Equity net payout (eqnpo_12m); Net total issuance (netis_at)
Distress and safety	Ohlson O-score (o_score); Altman Z-score (z_score); Quality minus Junk: Safety (qmj_safety)
Skewness	Coskewness (coskew_21d); Idiosyncratic skewness from the CAPM (iskew_capm_21d); Total skewness (rskew_21d)

Table 17 reports the estimates and inference underlying the figure. The standard error is the sample standard deviation of the category average across the 5,000 common accepted joint draws. The q-values apply Benjamini–Hochberg to the fixed family of 18 displayed

tilt–category averages.

Table 17. Anomaly contributions by behavioral tilt and economic category on the common 20% AUM scale

Tilt	Category	N	Estimate	SE	95% CI	BH q
Overconfidence	Market beta	3	202.6	72.7	[60.1, 345.1]	0.025
	Duration/Growth	3	108.3	47.9	[14.4, 202.2]	0.043
	Value	10	90.2	47.6	[-3.1, 183.5]	0.087
	Volatility	5	71.5	28.9	[14.9, 128.1]	0.032
	Investment	10	70.6	30.2	[11.4, 129.8]	0.039
	Momentum	6	33.9	41.7	[-47.8, 115.7]	0.468
	Post-news drift	3	9.9	16.4	[-22.1, 42.0]	0.543
	Turnover	2	-32.0	35.8	[-102.1, 38.1]	0.445
Lottery preference	Volatility	5	621.2	225.4	[179.4, 1063.0]	0.025
	Maximum returns	3	437.5	160.8	[122.4, 752.6]	0.025
	Issuance	3	421.9	150.9	[126.2, 717.7]	0.025
	Distress and safety	3	391.8	145.0	[107.7, 675.9]	0.025
	Value	10	243.2	94.7	[57.5, 428.8]	0.031
	Skewness	3	88.5	36.1	[17.7, 159.4]	0.032
	Market beta	3	64.2	64.8	[-62.8, 191.3]	0.414
Disposition effect	Momentum	6	115.2	58.4	[0.7, 229.7]	0.080
	Post-news drift	3	-20.3	28.9	[-77.0, 36.3]	0.510
Extrapolation	Value	10	54.0	41.5	[-27.5, 135.4]	0.269

D Realized Anomaly-Returns

Table 18 reports the complete equal-weighted realized-return evidence used for the 139-anomaly headline benchmark. Returns use formation-month assignments and next-month stock returns; a missing constituent return does not discard the other supported stocks in that anomaly-month leg.

Table 18. Realized equal-weighted anomaly returns

Feature	Anomaly	Mean (bp)	NW SE	p	BH q	Months	Support med./min. (L; S)
aliq_at	Liquidity of book assets	-50.3	427.2	0.906	0.947	119	23/18; 23/19
aliq_mat	Liquidity of market assets	463.3	524.0	0.377	0.557	119	23/20; 23/19
ami_126d	Amihud Measure	-583.7	449.9	0.195	0.360	119	26/22; 27/23
at_be	Book leverage	321.7	291.8	0.270	0.447	119	29/24; 29/25
at_gr1	Asset Growth	-275.7	413.9	0.505	0.651	119	28/25; 29/24
at_me	Assets-to-market	-1115.8	452.4	0.014	0.046	119	28/24; 29/24
at_turnover	Capital turnover	1225.6	393.5	0.002	0.009	119	29/24; 28/24
be_gr1a	Change in common equity	-778.2	387.7	0.045	0.109	119	28/24; 28/24
be_me	Book-to-market equity	-1126.2	489.1	0.021	0.065	119	28/24; 29/24
beta_60m	Market Beta	433.8	513.6	0.398	0.569	119	26/21; 26/20
beta_dimson_21d	Dimson beta	456.5	423.3	0.281	0.454	119	24/18; 24/20
betabab_1260d	Frazzini-Pedersen market beta	1413.9	591.5	0.017	0.053	119	20/16; 20/16
betadown_252d	Downside beta	150.9	510.7	0.768	0.861	119	26/22; 26/22
bev_mev	Book-to-market enterprise value	-1613.7	500.3	0.001	0.007	119	27/23; 28/23
bidaskhl_21d	The high-low bid-ask spread	-978.6	399.1	0.014	0.047	119	28/23; 28/22
capex_abn	Abnormal corporate investment	-318.2	431.7	0.461	0.613	119	20/16; 20/17
capx_gr1	CAPEX growth (1 year)	293.5	372.5	0.431	0.599	119	22/18; 22/18
capx_gr2	CAPEX growth (2 years)	-186.1	538.8	0.730	0.831	119	22/18; 22/18
capx_gr3	CAPEX growth (3 years)	-420.5	480.2	0.381	0.558	119	21/17; 21/18
cash_at	Cash-to-assets	-82.1	354.9	0.817	0.901	119	28/24; 28/25
coa_gr1a	Change in current operating assets	310.5	460.8	0.500	0.650	119	23/19; 23/19
col_gr1a	Change in current operating liabilities	73.3	411.6	0.859	0.925	119	23/20; 23/20
cop_at	Cash-based operating profits-to-book assets	1125.9	415.5	0.007	0.025	119	24/20; 23/19
cop_atl1	Cash-based operating profits-to-lagged book assets	840.3	429.3	0.050	0.118	119	23/20; 23/19
corr_1260d	Market correlation	-163.9	474.0	0.730	0.831	119	20/15; 21/18
coskew_21d	Coskewness	706.0	331.6	0.033	0.091	119	24/19; 23/19

Feature	Anomaly	Mean (bp)	NW SE	p	BH q	Months	Support med./min. (L; S)
cowc_gr1a	Change in current operating working capital	753.7	368.3	0.041	0.101	119	23/18; 24/19
dbnetis_at	Net debt issuance	965.5	289.5	< 0.001	0.005	119	28/25; 28/24
debt_gr3	Growth in book debt (3 years)	440.7	395.0	0.265	0.447	119	24/21; 24/20
debt_me	Debt-to-market	-1304.0	467.3	0.005	0.022	119	28/24; 28/23
dgp_dsale	Change gross margin minus change sales	-81.3	304.9	0.790	0.878	119	21/17; 21/16
div12m_me	Dividend yield	-371.7	412.6	0.368	0.556	119	17/9; 17/9
dolvol_126d	Dollar trading volume	-468.8	423.1	0.268	0.447	119	26/22; 27/22
dolvol_var_126d	Coefficient of variation for dollar trading volume	811.4	370.4	0.028	0.081	119	27/23; 26/22
dsale_dinv	Change sales minus change Inventory	698.1	505.3	0.167	0.323	119	19/16; 19/16
dsale_drec	Change sales minus change receivables	245.4	354.1	0.488	0.640	119	22/16; 22/19
dsale_dsga	Change sales minus change SG&A	-178.0	483.6	0.713	0.831	117	19/11; 19/12
earnings_variability	Earnings variability	737.0	377.5	0.051	0.118	119	25/19; 25/19
ebit_bev	Return on net operating assets	2018.7	333.9	< 0.001	< 0.001	119	28/24; 27/24
ebit_sale	Profit margin	528.2	426.4	0.215	0.384	119	28/25; 28/24
ebitda_mev	Ebitda-to-market enterprise value	535.7	417.9	0.200	0.361	119	29/24; 28/23
emp_gr1	Hiring rate	-14.6	459.7	0.975	0.991	119	22/12; 22/13
eq_dur	Equity duration	57.8	454.0	0.899	0.947	119	26/21; 26/20
eqnetis_at	Net equity issuance	1526.2	537.1	0.004	0.020	119	14/11; 14/10
eqnpo_12m	Equity net payout	904.9	407.2	0.026	0.076	119	29/25; 40/26
eqnpo_me	Net payout yield	973.5	515.6	0.059	0.132	119	9/6; 9/6
f_score	Pitroski F-score	377.2	402.8	0.349	0.545	119	30/21; 30/19
fcf_me	Free cash flow-to-price	1742.7	507.0	< 0.001	0.004	119	23/20; 22/18
fnl_gr1a	Change in financial liabilities	731.7	338.8	0.031	0.086	119	28/24; 28/24
gp_at	Gross profits-to-assets	1779.4	468.5	< 0.001	0.001	119	23/20; 23/19
gp_atl1	Gross profits-to-lagged assets	1854.7	539.3	< 0.001	0.004	119	23/19; 23/18
inv_gr1	Inventory growth	201.1	523.2	0.701	0.830	119	19/16; 20/16
inv_gr1a	Inventory change	405.6	391.2	0.300	0.479	119	27/20; 27/20
iskew_capm_21d	Idiosyncratic skewness from the CAPM	-25.1	366.5	0.945	0.981	119	23/20; 24/21
ival_me	Intrinsic value-to-market	-392.9	535.6	0.463	0.613	119	16/10; 16/11
ivol_capm_21d	Idiosyncratic volatility from the CAPM (21 days)	1173.6	423.0	0.006	0.022	119	24/20; 23/20
ivol_capm_252d	Idiosyncratic volatility from the CAPM (252 days)	604.8	385.1	0.116	0.238	119	27/23; 26/21
kz_index	Kaplan-Zingales index	-636.3	431.9	0.141	0.276	119	14/11; 14/10

Feature	Anomaly	Mean (bp)	NW SE	p	BH q	Months	Support med./min. (L; S)
lnoa_gr1a	Change in long-term net operating assets	282.4	383.0	0.461	0.613	119	27/23; 26/23
lti_gr1a	Change in long-term investments	-331.6	366.5	0.366	0.556	119	24/19; 28/19
market_equity	Market Equity	-926.2	335.8	0.006	0.022	119	28/23; 29/25
mispricing_mgmt	Mispricing factor: Management	345.2	411.1	0.401	0.569	119	29/25; 29/24
mispricing_perf	Mispricing factor: Performance	2368.1	455.2	< 0.001	< 0.001	119	29/25; 28/24
ncoa_gr1a	Change in noncurrent operating assets	-83.9	389.7	0.829	0.908	119	23/20; 23/19
ncol_gr1a	Change in noncurrent operating liabilities	-746.8	439.4	0.089	0.188	119	23/20; 23/19
netdebt_me	Net debt-to-price	1264.8	425.0	0.003	0.014	119	28/24; 28/24
netis_at	Net total issuance	1068.9	465.2	0.022	0.065	119	14/11; 14/10
nfna_gr1a	Change in net financial assets	837.3	372.5	0.025	0.073	119	28/24; 28/24
ni_ar1	Earnings persistence	-424.2	392.9	0.280	0.454	119	25/20; 25/18
ni_be	Return on equity	1609.2	392.2	< 0.001	< 0.001	119	29/25; 28/24
ni_inc8q	Number of consecutive quarters with earnings increases	-392.5	393.8	0.319	0.504	100	29/22; 38/24
ni_ivol	Earnings volatility	-458.2	351.8	0.193	0.360	119	25/20; 25/20
ni_me	Earnings-to-price	1107.9	394.7	0.005	0.021	119	29/24; 28/24
niq_at	Quarterly return on assets	1454.7	331.5	< 0.001	< 0.001	119	29/25; 28/23
niq_at_chg1	Change in quarterly return on assets	679.3	356.7	0.057	0.130	119	28/23; 28/23
niq_be	Quarterly return on equity	1205.8	344.0	< 0.001	0.004	119	29/25; 28/24
niq_be_chg1	Change in quarterly return on equity	503.1	375.3	0.180	0.343	119	28/24; 28/23
niq_su	Standardized earnings surprise	194.3	410.3	0.636	0.789	117	27/23; 27/23
nncoa_gr1a	Change in net noncurrent operating assets	-70.7	385.1	0.854	0.925	119	23/19; 23/19
noa_at	Net operating assets	1674.8	480.2	< 0.001	0.004	119	23/19; 23/19
noa_gr1a	Change in net operating assets	380.6	428.7	0.375	0.557	119	23/18; 23/19
o_score	Ohlson O-score	2172.3	480.4	< 0.001	< 0.001	119	24/19; 23/19
oaccruals_at	Operating accruals	-358.9	415.4	0.388	0.561	119	23/20; 24/19
oaccruals_ni	Percent operating accruals	-14.8	503.1	0.977	0.991	119	23/19; 23/19
ocf_at	Operating cash flow to assets	1113.6	415.2	0.007	0.027	119	29/25; 28/25
ocf_at_chg1	Change in operating cash flow to assets	363.2	315.8	0.250	0.433	119	28/24; 28/24
ocf_me	Operating cash flow-to-market	-189.3	499.3	0.705	0.830	119	28/25; 28/24
ocfq_saleq_std	Cash flow volatility	1157.5	480.6	0.016	0.052	119	22/19; 21/17
op_at	Operating profits-to-book assets	1232.5	336.1	< 0.001	0.002	119	29/25; 28/24

Feature	Anomaly	Mean (bp)	NW SE	p	BH q	Months	Support med./min. (L; S)
op_atl1	Operating profits-to-lagged book assets	1092.5	389.1	0.005	0.021	119	29/25; 28/25
ope_be	Operating profits-to-book equity	1150.7	383.7	0.003	0.013	119	29/25; 28/24
ope_bel1	Operating profits-to-lagged book equity	982.1	396.1	0.013	0.046	119	28/24; 27/24
opex_at	Operating leverage	1771.2	593.5	0.003	0.014	119	23/20; 23/19
pi_nix	Taxable income-to-book income	309.2	398.9	0.438	0.603	119	21/16; 21/15
ppeinv_gr1a	Change PPE and Inventory	463.0	404.3	0.252	0.433	119	24/18; 24/18
prc	Price per share	-943.2	477.4	0.048	0.116	119	28/24; 29/25
prc_highprc_252d	Current price to high price over last year	2144.3	625.7	< 0.001	0.004	119	26/22; 26/21
qmj	Quality minus Junk: Composite	1936.0	434.0	< 0.001	< 0.001	119	25/21; 25/18
qmj_growth	Quality minus Junk: Growth	1463.4	336.8	< 0.001	< 0.001	119	25/21; 25/20
qmj_prof	Quality minus Junk: Profitability	1217.2	362.1	< 0.001	0.005	119	29/25; 28/24
qmj_safety	Quality minus Junk: Safety	2334.2	537.2	< 0.001	< 0.001	119	29/25; 28/25
rd_me	R&D-to-market	1407.1	824.2	0.088	0.188	117	8/6; 8/6
rd_sale	R&D-to-sales	955.8	627.4	0.128	0.253	117	8/6; 8/6
ret_12_1	Price momentum t-12 to t-1	2346.0	497.6	< 0.001	< 0.001	119	28/24; 28/24
ret_12_7	Price momentum t-12 to t-7	1835.2	406.6	< 0.001	< 0.001	119	29/25; 28/24
ret_1_0	Short-term reversal	-267.0	504.5	0.597	0.754	119	28/24; 29/24
ret_3_1	Price momentum t-3 to t-1	2021.3	489.3	< 0.001	< 0.001	119	28/24; 28/24
ret_60_12	Long-term reversal	-591.9	462.1	0.200	0.361	119	22/16; 23/18
ret_6_1	Price momentum t-6 to t-1	1971.5	484.4	< 0.001	< 0.001	119	29/25; 28/24
ret_9_1	Price momentum t-9 to t-1	2288.9	511.3	< 0.001	< 0.001	119	28/24; 28/24
rmax1_21d	Maximum daily return	774.5	430.0	0.072	0.158	119	24/20; 23/20
rmax5_21d	Highest 5 days of return	956.7	452.1	0.034	0.092	119	24/20; 23/19
rmax5_rvol_21d	Highest 5 days of return scaled by volatility	182.9	456.1	0.688	0.830	119	23/19; 23/19
rskew_21d	Total skewness	-185.2	406.1	0.648	0.797	119	23/19; 24/20
rvol_21d	Return volatility	1326.9	503.9	0.008	0.030	119	24/20; 23/20
sale_bev	Asset turnover	1419.8	449.9	0.002	0.009	119	28/24; 28/23
sale_emp_gr1	Labor force efficiency	-566.0	470.8	0.229	0.403	119	22/13; 22/13
sale_gr1	Sales Growth (1 year)	-3.2	313.3	0.992	0.999	119	28/24; 28/24
sale_gr3	Sales Growth (3 years)	-557.4	332.0	0.093	0.193	119	27/22; 27/22
sale_me	Sales-to-market	-124.1	404.9	0.759	0.858	119	28/24; 28/24
saleq_gr1	Sales growth (1 quarter)	-294.8	355.3	0.407	0.571	119	22/19; 22/18
saleq_su	Standardized Revenue surprise	-53.7	428.1	0.900	0.947	117	21/18; 21/17
seas_1_1an	Year 1-lagged return, annual	906.0	437.7	0.038	0.097	119	28/24; 28/24

Feature	Anomaly	Mean (bp)	NW SE	p	BH q	Months	Support med./min. (L; S)
seas_1_1na	Year 1-lagged return, nonannual	1962.2	522.1	< 0.001	0.002	119	28/23; 28/24
seas_2_5an	Years 2-5 lagged returns, annual	327.8	364.1	0.368	0.556	119	22/17; 22/15
seas_2_5na	Years 2-5 lagged returns, nonannual	-394.6	531.0	0.457	0.613	119	21/16; 21/17
seas_6_10an	Years 6-10 lagged returns, annual	14.0	464.4	0.976	0.991	119	14/7; 14/7
seas_6_10na	Years 6-10 lagged returns, nonannual	328.1	634.6	0.605	0.758	119	13/7; 13/6
sti_gr1a	Change in short-term investments	585.8	338.9	0.084	0.182	98	51/21; 26/22
taccruals_at	Total accruals	-825.6	393.8	0.036	0.094	119	23/19; 23/20
taccruals_ni	Percent total accruals	-236.0	390.7	0.546	0.696	119	24/19; 23/19
tangibility	Asset tangibility	-175.4	486.8	0.719	0.831	119	23/19; 23/18
tax_gr1a	Tax expense surprise	-0.6	372.5	0.999	0.999	119	28/24; 28/24
turnover_126d	Share turnover	-218.8	556.4	0.694	0.830	119	27/22; 26/22
turnover_var_126d	Coefficient of variation for share turnover	786.5	376.2	0.037	0.094	119	27/23; 26/22
z_score	Altman Z-score	1970.3	539.3	< 0.001	0.002	119	23/19; 23/20
zero_trades_126d	Number of zero trades with turnover as tiebreaker (6 months)	-218.8	556.4	0.694	0.830	119	27/22; 26/22
zero_trades_21d	Number of zero trades with turnover as tiebreaker (1 month)	-813.6	525.6	0.122	0.245	119	24/20; 23/20
zero_trades_252d	Number of zero trades with turnover as tiebreaker (12 months)	64.6	537.9	0.904	0.947	119	27/22; 26/22

E Multiple Testing

We use the Benjamini–Hochberg procedure to control the false discovery rate within the fixed creator family (Benjamini and Hochberg, 1995). Table 19 reports the global adjustment across all 91 creator cells and sensitivity adjustments within each tilt. The count of 23 discoveries uses the global family and a 5% false-discovery-rate threshold.

Table 19. Multiple-testing audit for creator predictions

Hypothesis family	Tests	Raw $p < 0.05$	BH $q \leq 0.10$	BH $q \leq 0.05$	Minimum BH q
Creator cells, global	91	46	46	23	0.030
Creator cells, disposition	9	5	5	0	0.066
Creator cells, lottery	30	22	22	21	0.012
Creator cells, overconfidence	42	15	13	0	0.072
Creator cells, extrapolation	10	4	4	0	0.051

F Anomaly Results Without AUM Normalization

The four counterfactuals move different average shares of household AUM because the bias measures differ in coverage and bin structure. Equation (44) normalizes each contribution to a reallocation of 20% of average household AUM so that the four families are comparable. This section reports the corresponding estimates without that normalization. The signs, test statistics, p-values, q-values, and qualitative conclusions are unchanged.

Table 20. Pooled tests of creator predictions (without AUM normalization)

Behavioral tilt	Positive sign	Average (bp)	Bootstrap SE	One-sided p
Disposition effect	6/9	81.5	47.2	0.042
Lottery preference	25/30	322.1	115.2	0.003
Overconfidence	36/42	119.2	41.7	0.002
Extrapolation	6/10	21.2	16.3	0.097
Equal tilt average	73/91	136.0	45.0	0.001

Notes: The table reports annualized equal-weighted contributions for the creator predictions. Cells are equally weighted within each tilt, and the headline equally weights the four tilt averages. Positive estimates agree with the prediction. Standard errors use exactly 5,000 common joint bootstrap draws; one-sided p-values use the standard-normal distribution.

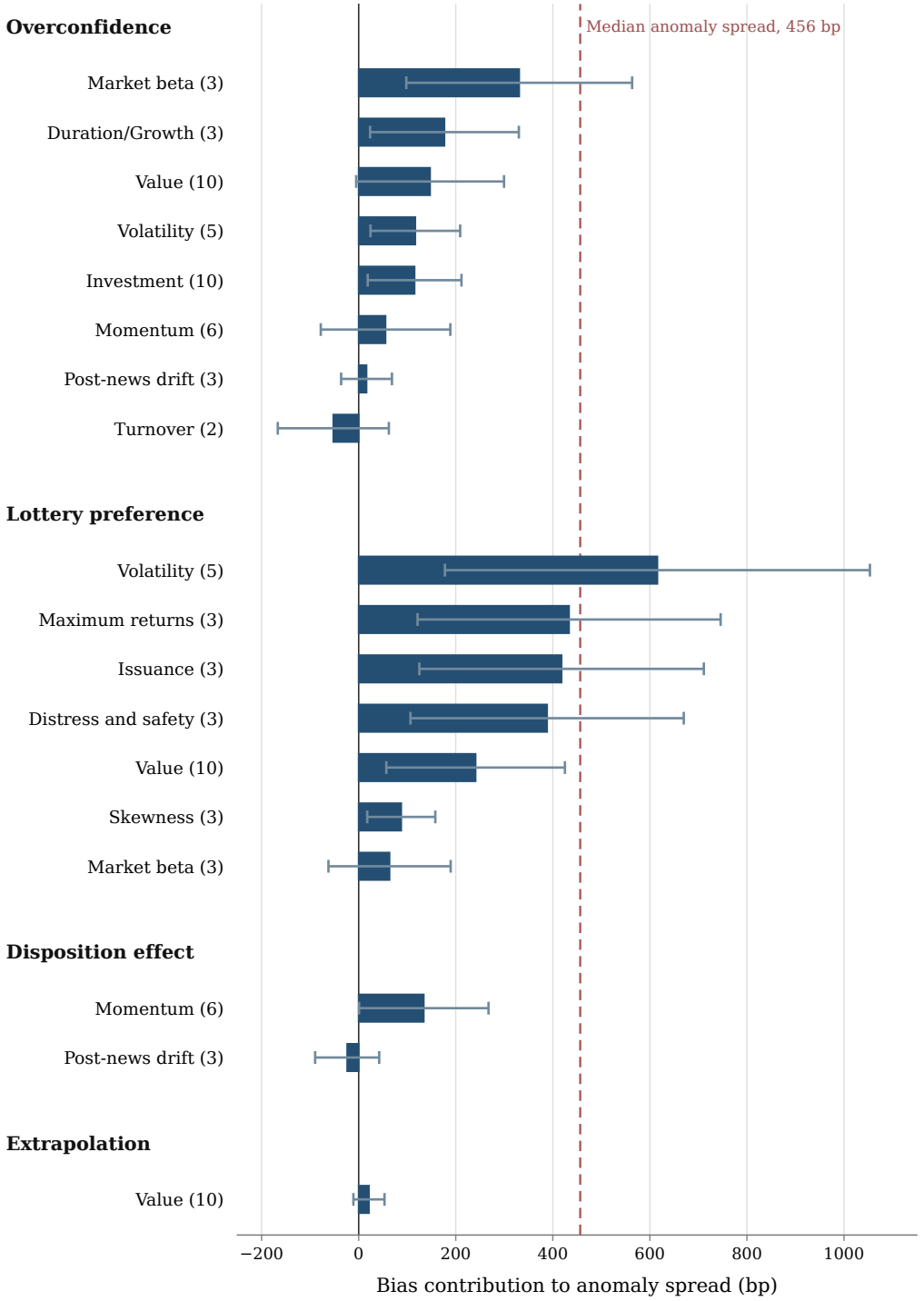


Figure 7. Anomaly contributions by behavioral tilt and economic category without AUM normalization. Each bar is the simple average of the annualized equal-weighted contributions in the category. Whiskers report 95% normal-bootstrap confidence intervals from 5,000 common joint draws.

Table 21. Anomaly contributions by behavioral tilt and economic category without AUM normalization

Tilt	Category	N	Estimate	SE	95% CI	BH q
Overconfidence	Market beta	3	330.8	118.7	[98.1, 563.5]	0.025
	Duration/Growth	3	176.8	78.2	[23.5, 330.0]	0.043
	Value	10	147.2	77.7	[-5.1, 299.5]	0.087
	Volatility	5	116.8	47.1	[24.4, 209.2]	0.032
	Investment	10	115.2	49.3	[18.6, 211.8]	0.039
	Momentum	6	55.4	68.1	[-78.1, 188.8]	0.468
	Post-news drift	3	16.2	26.7	[-36.1, 68.6]	0.543
	Turnover	2	-52.2	58.4	[-166.6, 62.2]	0.445
Lottery preference	Volatility	5	615.6	223.4	[177.8, 1053.4]	0.025
	Maximum returns	3	433.6	159.3	[121.3, 745.8]	0.025
	Issuance	3	418.1	149.5	[125.1, 711.2]	0.025
	Distress and safety	3	388.3	143.6	[106.8, 669.9]	0.025
	Value	10	241.0	93.9	[57.0, 425.0]	0.031
	Skewness	3	87.7	35.8	[17.6, 157.9]	0.032
	Market beta	3	63.6	64.2	[-62.3, 189.5]	0.414
Disposition effect	Momentum	6	134.2	68.0	[0.8, 267.5]	0.080
	Post-news drift	3	-23.7	33.7	[-89.7, 42.3]	0.510
Extrapolation	Value	10	21.2	16.3	[-10.8, 53.3]	0.269

G Model Details Behind the Equilibrium Objects

This appendix fills in the derivation behind the equilibrium objects introduced in Section 3.1. The main text defines the pass-through matrix, the total-elasticity matrix, and the flow multiplier. Here we show how those objects arise from standard log-linearizations of returns, demand, and risk.

G.1 Campbell–Shiller and Pass-Through

For stock-level log returns, the Campbell–Shiller log-linearization gives

$$\boldsymbol{r}_{t+1} = \zeta \mathbf{1} + (1-\rho)\boldsymbol{d}_{t+1} + \rho \boldsymbol{p}_{t+1} - \boldsymbol{p}_t, \tag{45}$$

where ρ is the usual price-to-(price+dividend) linearization constant (Cochrane, 2008). In the baseline specification, we use the monthly conversion of the annual approximation point in Cochrane (2008), so the monthly dividend-to-(price+dividend) ratio linearization constant is about 0.3%.

The main text allows the conditional price law to be a general linear function of current log-prices, as in (10). Taking conditional expectations in (45) and substituting (10) yields the expected-return mapping in (11), with

$$\bar{\boldsymbol{\mu}}_t \equiv \zeta \mathbf{1} + (1-\rho)\mathbb{E}_t[\boldsymbol{d}_{t+1}] + \rho \boldsymbol{c}_t.$$

Thus $\boldsymbol{\Theta}_t = \boldsymbol{I} - \rho \boldsymbol{\Psi}_t$ is the price pass-through to expected returns.

G.2 Log-Linearized Demand and Total Elasticity

Let $\boldsymbol{\beta}_t \in \mathbb{R}^{N_t \times K}$ denote stock-level factor loadings on the K orthogonal factors defined in Section 3.4, let $\boldsymbol{\beta}_{k,t} \in \mathbb{R}^{N_t}$ denote the k-th column of $\boldsymbol{\beta}_t$, let $\boldsymbol{v}_t \in \mathbb{R}^K$ denote the variances of those K factors, and let $\boldsymbol{\sigma}_t^2 \in \mathbb{R}^{N_t}$ denote stock-level idiosyncratic variances. Then each $\boldsymbol{\beta}_{k,t} v_{k,t} \in \mathbb{R}^{N_t}$ is the stock-level covariance vector with factor k, and the full systematic-risk component is $\sum_{k=1}^{K} \boldsymbol{\beta}_{k,t} v_{k,t} = \boldsymbol{\beta}_t \boldsymbol{v}_t$.

Standard CARA or CRRA demand is a function of next-period expected returns and covariances. Following Koijen and Yogo (2019), we also allow an investor-specific latent demand shifter. While the demand system below can be derived with standard expected utility preferences, it is a general demand function that captures many types of preferences and motives. In the log-linearized system with latent demand, there are four important ingredients: expected returns $\boldsymbol{\mu}_t$, the K factor-specific systematic-risk components $\{\boldsymbol{\beta}_{k,t} v_{k,t}\}_{k=1}^{K}$, idiosyncratic risk $\boldsymbol{\sigma}_t^2$, and latent demand $\boldsymbol{\xi}_{i,t}$. We follow Davis et al. (2026) by taking a log-linearization of demand for each investor i around these four objects:

$$\boldsymbol{q}_{i,t} = a_{i,t}\mathbf{1} + \bar{\boldsymbol{q}}_{i,t} + \boldsymbol{C}_{i,t}\,\boldsymbol{\mu}_t - \sum_{k=1}^{K} \boldsymbol{D}_{i,k,t}\boldsymbol{\beta}_{k,t}\boldsymbol{v}_{k,t} - \boldsymbol{E}_{i,t}\boldsymbol{\sigma}_t^2 + \boldsymbol{\xi}_{i,t}, \tag{46}$$

where $\bar{\boldsymbol{q}}_{i,t} \in \mathbb{R}^{N_t}$ collects investor-specific intercept terms beyond the AUM intercept component $(a_{i,t})$ that do not load on current prices, $\boldsymbol{C}_{i,t} \in \mathbb{R}^{N_t \times N_t}$ is the coefficient matrix for expected returns, the collection $\{\boldsymbol{D}_{i,k,t}\}_{k=1}^{K}$ with each $\boldsymbol{D}_{i,k,t} \in \mathbb{R}^{N_t \times N_t}$ captures preferences over factor-k systematic risk, $\boldsymbol{E}_{i,t} \in \mathbb{R}^{N_t \times N_t}$ captures preferences over idiosyncratic risk, and $\boldsymbol{\xi}_{i,t} \in \mathbb{R}^{N_t}$ is the usual latent demand shifter. The accounting identity in (9) is what makes the AUM term enter one-for-one in log-shares demanded.

Future risk is not exogenous to prices. Empirically, lower prices are associated with higher future volatility and stronger future comovement in returns (Black, 1976; Ang and Chen, 2002; Dew-Becker, Giglio, and Molavi, 2025). He et al. (2025) discuss these effects in demand-system asset-pricing models. To capture this, we allow:

$$\boldsymbol{v}_t = \bar{\boldsymbol{v}}_t + \boldsymbol{\Gamma}_t^{V}\boldsymbol{p}_t, \qquad \boldsymbol{\sigma}_t^2 = \bar{\boldsymbol{\sigma}}_t^2 + \boldsymbol{\Gamma}_t^{\sigma}\boldsymbol{p}_t, \tag{47}$$

where $\boldsymbol{\Gamma}_t^{V} \in \mathbb{R}^{K \times N_t}$ captures the sensitivity of factor variances to price changes, $\boldsymbol{\Gamma}_t^{\sigma} \in \mathbb{R}^{N_t \times N_t}$ captures the sensitivity of stock-level idiosyncratic variances to current prices, and $\bar{\boldsymbol{v}}_t$ and $\bar{\boldsymbol{\sigma}}_t^2$ are intercept terms. Let $\bar{v}_{k,t}$ denote the k-th element of $\bar{\boldsymbol{v}}_t$ and let $\boldsymbol{\gamma}_{k,t}^{V\prime}$ denote the k-th

row of $\mathbf{\Gamma}_t^V$. Combining (46), (11), and (47) gives

$$\begin{aligned}\boldsymbol{q}_{i,t} &= a_{i,t}\mathbf{1} + \bar{\boldsymbol{q}}_{i,t} + \boldsymbol{C}_{i,t}\big(\bar{\boldsymbol{\mu}}_t - \boldsymbol{\Theta}_t \boldsymbol{p}_t\big) \\ &\quad - \sum_{k=1}^{K} \boldsymbol{D}_{i,k,t}\boldsymbol{\beta}_{k,t}\big(\bar{v}_{k,t} + \boldsymbol{\gamma}_{k,t}^{V\prime}\boldsymbol{p}_t\big) - \boldsymbol{E}_{i,t}\big(\bar{\boldsymbol{\sigma}}_t^2 + \mathbf{\Gamma}_t^{\sigma}\boldsymbol{p}_t\big) + \boldsymbol{\xi}_{i,t} \\ &= \boldsymbol{\kappa}_{i,t} - \boldsymbol{\eta}_{i,t}^{\text{tot}}\boldsymbol{p}_t, \end{aligned} \tag{48}$$

where

$$\begin{aligned}\boldsymbol{\kappa}_{i,t} &\equiv a_{i,t}\mathbf{1} + \bar{\boldsymbol{q}}_{i,t} + \boldsymbol{C}_{i,t}\bar{\boldsymbol{\mu}}_t - \sum_{k=1}^{K} \boldsymbol{D}_{i,k,t}\boldsymbol{\beta}_{k,t}\bar{v}_{k,t} - \boldsymbol{E}_{i,t}\bar{\boldsymbol{\sigma}}_t^2 + \boldsymbol{\xi}_{i,t}, \\ \boldsymbol{\eta}_{i,t}^{\text{tot}} &\equiv \boldsymbol{C}_{i,t}\boldsymbol{\Theta}_t + \sum_{k=1}^{K} \boldsymbol{D}_{i,k,t}\boldsymbol{\beta}_{k,t}\boldsymbol{\gamma}_{k,t}^{V\prime} + \boldsymbol{E}_{i,t}\mathbf{\Gamma}_t^{\sigma}. \end{aligned}$$

This is the reduced-form total-elasticity representation used in (12). The first term is the expected-return channel, the second collects factor-by-factor risk channels, and the third is the idiosyncratic-risk channel.

H Proofs of Methodological Propositions

In this appendix, we prove the non-obvious methodological claims used in the main text. Our goal is to keep the main text readable while making the mapping from the structural assumptions to the empirical objects fully transparent.

Proof of Proposition 1. Start from the estimated demand equation

$$\boldsymbol{q}_{i,t} = \boldsymbol{\kappa}_{i,t} - \boldsymbol{\eta}_{i,t}^{\text{tot}}\boldsymbol{p}_t.$$

Consider a counterfactual perturbation that changes investor i's log AUM by $\Delta^{\text{cf}}a_{i,t}$ while holding $\boldsymbol{\eta}_{i,t}^{\text{tot}}$ fixed. The counterfactual change in log demand is therefore

$$\Delta^{\text{cf}}\boldsymbol{q}_{i,t} = \Delta^{\text{cf}}a_{i,t}\mathbf{1} - \boldsymbol{\eta}_{i,t}^{\text{tot}}\Delta^{\text{cf}}\boldsymbol{p}_t.$$

Applying the log-linearized market-clearing condition,

$$\mathbf{0} = \sum_i \bar{\boldsymbol{S}}_{i,t} \Delta^{\mathrm{cf}} \boldsymbol{q}_{i,t},$$

gives

$$\mathbf{0} = \Delta^{\mathrm{cf}} \boldsymbol{f}_t - \sum_i \bar{\boldsymbol{S}}_{i,t} \boldsymbol{\eta}_{i,t}^{\mathrm{tot}} \Delta^{\mathrm{cf}} \boldsymbol{p}_t.$$

Rearranging yields

$$\left(\sum_i \bar{\boldsymbol{S}}_{i,t} \boldsymbol{\eta}_{i,t}^{\mathrm{tot}} \right) \Delta^{\mathrm{cf}} \boldsymbol{p}_t = \Delta^{\mathrm{cf}} \boldsymbol{f}_t.$$

If the aggregate elasticity matrix is invertible, left-multiplication by its inverse gives (15). □

Variance-decomposition note. The Standalone-variance shares in (30) are normalized standalone variance contributions and are nonnegative by construction. The R^2 Shapley shares are Shapley allocations of the fit-improvement value function in (34). Because that value function is nonnegative and monotone, the resulting monthly shares are also nonnegative and sum to one within each partition. The signed Shapley shares in (36) additionally allocate each covariance term equally between its two investor groups. These shares sum to one and can be negative. The main analysis reports all three allocations.

I Additional Details

In this appendix, we collect the algebra behind the low-rank basis construction and the corresponding regression representation used in Section 3.4.

I.1 Characteristics, Portfolio Weights, and Loadings

Define the rank-transformed characteristic panel by

$$\mathcal{R}(c_{n,k,t}) \equiv \frac{\mathrm{rank}(c_{n,k,t}) - 0.5}{N_t} - 0.5, \qquad Z_{n,k,t} \equiv \mathcal{R}(c_{n,k,t}), \tag{49}$$

and let $\boldsymbol{Z}_t \in \mathbb{R}^{N_t \times K}$ stack these transformed characteristics. The orthonormalized characteristic basis is

$$\widetilde{\boldsymbol{X}}_t \equiv \boldsymbol{Z}_t(\boldsymbol{Z}_t'\boldsymbol{Z}_t)^{-1/2}, \qquad \widetilde{\boldsymbol{X}}_t'\widetilde{\boldsymbol{X}}_t = \boldsymbol{I}_K. \tag{50}$$

We keep the full orthonormalized characteristic span and do not further reduce its dimension. Any orthogonal rotation inside that span is therefore innocuous. Write

$$\boldsymbol{X}_t \equiv \widetilde{\boldsymbol{X}}_t\boldsymbol{U}, \qquad \boldsymbol{U}'\boldsymbol{U} = \boldsymbol{I}_K, \qquad \boldsymbol{X}_t'\boldsymbol{X}_t = \boldsymbol{I}_K. \tag{51}$$

for an orthogonal matrix $\boldsymbol{U}$. So the column space of $\boldsymbol{X}_t$ is exactly the same as the column space of the original ranked-characteristic matrix $\boldsymbol{Z}_t$, and orthonormality is preserved because

$$\boldsymbol{X}_t'\boldsymbol{X}_t = \boldsymbol{U}'\widetilde{\boldsymbol{X}}_t'\widetilde{\boldsymbol{X}}_t\boldsymbol{U} = \boldsymbol{U}'\boldsymbol{U} = \boldsymbol{I}_K.$$

The associated factor returns are

$$\boldsymbol{r}_{t+1}^F = \boldsymbol{X}_t'\boldsymbol{r}_{t+1}. \tag{52}$$

The fitted return component in the characteristic span is therefore

$$\widehat{\boldsymbol{r}}_{t+1} = \boldsymbol{X}_t\boldsymbol{r}_{t+1}^F = \boldsymbol{X}_t\boldsymbol{X}_t'\boldsymbol{r}_{t+1}. \tag{53}$$

Equation (53) makes the triple role of $\boldsymbol{X}_t$ explicit. Its columns are portfolio-weight vectors because they form the characteristic-managed portfolios whose returns are $\boldsymbol{r}_{t+1}^F$. The same columns are stock-level factor loadings because they map those factor returns back into stock space. And because $\boldsymbol{X}_t$ is just an orthonormal rotation of the original ranked characteristics, it is also an orthonormal set of asset characteristics.

The accounting identity in the main text,

$$\boldsymbol{q}_{i,t} = a_{i,t}\mathbf{1} + \boldsymbol{w}_{i,t} - \boldsymbol{p}_t,$$

implies that a weight regression and a holdings regression differ only by the direct price

term. This is why it is convenient to estimate low-rank structure in portfolio weights and then translate it back into holdings elasticities.

I.2 Low-Rank Regression Form, Covariance Intuition, and Instruments

For an investor i in group $\mathcal{G}(i)$, start from the estimated holdings equation

$$\boldsymbol{q}_{i,t} = \boldsymbol{\kappa}_{i,t} - \boldsymbol{\eta}^{\text{tot}}_{\mathcal{G}(i),t}\boldsymbol{p}_t$$

and impose the low-rank elasticity restriction

$$\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G}(i),t} = (1 - \eta_{0,\mathcal{G}(i)})\boldsymbol{I} + \boldsymbol{X}_t\boldsymbol{\Gamma}_{\mathcal{G}(i)}\boldsymbol{X}_t'.$$

Using $\boldsymbol{q}_{i,t} = a_{i,t}\mathbf{1} + \boldsymbol{w}_{i,t} - \boldsymbol{p}_t$ gives

$$\boldsymbol{w}_{i,t} = (\boldsymbol{\kappa}_{i,t} - a_{i,t}\mathbf{1}) + \eta_{0,\mathcal{G}(i)}\boldsymbol{p}_t - \boldsymbol{X}_t\boldsymbol{\Gamma}_{\mathcal{G}(i)}\boldsymbol{X}_t'\boldsymbol{p}_t.$$

Absorbing the book-equity term $\boldsymbol{b}_t$ into the intercept and writing $\boldsymbol{g}_t = \boldsymbol{p}_t - \boldsymbol{b}_t$ yields the regression form in (19). Row n loads on the factor shock $\boldsymbol{X}_t'\boldsymbol{g}_t$ according to $\boldsymbol{x}_{n,t}'\boldsymbol{\Gamma}_{\mathcal{G}(i)}$. The coefficients $\eta_{0,\mathcal{G}(i)}$ and $\boldsymbol{\Gamma}_{\mathcal{G}(i)}$ are common within investor group, while the investor-month intercepts and the characteristic slopes are investor specific.

The same low-rank logic applies to pass-through. We estimate the date-t valuation gap's predictive loading for the realized 12-month return as

$$\mathbb{E}_t[\boldsymbol{R}^{(12)}_{t\to t+12}] = \boldsymbol{\alpha} + \gamma_t\mathbf{1} + \boldsymbol{C}_t\boldsymbol{\delta} + \left(\beta_0^{(12)}\boldsymbol{I} + \boldsymbol{X}_t\boldsymbol{\Gamma}_R^{(12)}\boldsymbol{X}_t'\right)\boldsymbol{g}_t.$$

Equation (21) maps this annual-horizon predictive response into the monthly pass-through matrix, and $\boldsymbol{\Psi}_t = (\boldsymbol{I} - \boldsymbol{\Theta}_t)/\rho$ gives the corresponding monthly price law. The scalar terms are direct own-stock responses, while the interaction terms are spillovers through the factor portfolios.

Finally, suppose $\hat{g}_{in,t}$ is a valid excluded instrument for $g_{n,t}$ in the investor-level panel

and define

$$\hat{z}_{i,l,t} \equiv \sum_{m=1}^{N_t} x_{m,l,t} \hat{g}_{im,t}.$$

Because $\boldsymbol{X}_t$ and $x_{n,k,t}$ are predetermined at date t, the interacted terms $x_{n,k,t}\hat{z}_{i,l,t}$ are valid instruments for the low-rank spillover regressors whenever the same exclusion restriction holds after interaction. In moment-condition form, the required conditions are

$$\begin{aligned} &\mathbb{E}[\hat{g}_{in,t} u_{i,n,t}] = 0, && \mathbb{E}[x_{n,k,t}\hat{z}_{i,l,t} u_{i,n,t}] = 0, \\ &\mathbb{E}[\hat{g}_{n,t} \nu^{(12)}_{n,t}] = 0, && \mathbb{E}[x_{n,k,t}\hat{g}^{\text{fac}}_{l,t} \nu^{(12)}_{n,t}] = 0 \qquad \text{for all } k, l. \end{aligned}$$

where the second line uses the stock-level pass-through instrument. So the interaction instruments do not create a new identifying assumption; they reuse the base instrument along additional predetermined directions in characteristic space.

I.3 Instrument, IV Panels, and Shrinkage

The monthly stock panel carries the sticky-mandate IV objects defined in the main text. On the demand side, estimation proceeds on the investor-stock-month panel, both for the broad demand system and for the four bias-specific systems. The dependent variable is log portfolio weight. The exogenous part of the residualization absorbs investor-month effects and investor-specific characteristic slopes. The endogenous regressors are the valuation gap $g_{n,t}$ and the sixteen low-rank spillover terms $x_{n,k,t}z_{l,t}$. The excluded instruments are the investor-level leave-out-own valuation gap $\hat{g}_{in,t}$ and the corresponding interacted terms $x_{n,k,t}\hat{z}_{i,l,t}$.

The pass-through panel uses the realized compounded stock return from $t+1$ through $t+12$. We require all twelve monthly returns to be present and consecutive rather than filling gaps. After removing stock and month fixed effects and controlling for date-t book equity, CAPM beta, ESG greenness, and ROE, the endogenous regressors are again $g_{n,t}$ and the sixteen low-rank spillover terms, and the excluded instruments are the stock-level valuation gap $\hat{g}_{n,t}$ and its interacted analogues. We then divide the fitted 12-month return-response block by -12 to express $\boldsymbol{\Theta}_t$ in monthly units, as in (21).

Our shrinkage step is empirical Bayes. For each unrestricted spillover entry, the posterior mean is

$$\tilde{\gamma}_j = \frac{\tau^2}{\tau^2 + v_j}\hat{\gamma}_j,$$

with v_j taken from the estimated covariance matrix and τ^2 chosen by marginal likelihood. We apply this rule jointly across the broad demand system and the four bias systems on the demand side, and separately on the pass-through side.

I.4 Common and Granular Components of Top-Tail Bias Tilts

This subsection describes the decomposition reported in Table 22. The object is the top-tail active tilt within each bias family. For overconfidence, the top tail is the highest tercile in each wealth decile. For lottery preference, the disposition effect, and extrapolation, the top tail is the top two quintiles in each wealth decile. Let s index these top-tail sleeves, let $w(s)$ denote sleeve s's wealth decile, and let $\boldsymbol{w}_{s,t}$ denote sleeve s's portfolio weights in month t. Let $\boldsymbol{w}^{\text{hh}}_{w(s),t}$ be the AUM-weighted portfolio of all non-residual household sleeves in the same wealth decile. The sleeve-level active tilt is

$$\boldsymbol{a}_{s,t} \equiv \boldsymbol{w}_{s,t} - \boldsymbol{w}^{\text{hh}}_{w(s),t}.$$

By construction, this active tilt compares a top-tail sleeve to households with the same wealth rank.

For each bias family b and month t, stack the sleeve tilts as columns of $\boldsymbol{A}^b_t$. We take the first left singular vector of $\boldsymbol{A}^b_t$, denoted $\boldsymbol{u}^b_{1,t}$, as the common cross-wealth behavioral direction. The common and granular components of sleeve s are

$$\boldsymbol{a}^{\text{com}}_{s,t} = \boldsymbol{u}^b_{1,t}(\boldsymbol{u}^b_{1,t})'\boldsymbol{a}_{s,t}, \qquad \boldsymbol{a}^{\text{gran}}_{s,t} = \boldsymbol{a}_{s,t} - \boldsymbol{a}^{\text{com}}_{s,t}.$$

The PC1 share in Table 22 is the share of $\|\boldsymbol{A}^b_t\|^2_F$ explained by this first component.

We then ask how much each component matters for prices. Let $\omega_{s,t}$ be sleeve s's AUM divided by the aggregate market capitalization of the demand-system universe in month t.

The stock-level flow pressures from the common and granular components are

$$\Delta \boldsymbol{f}_{b,t}^{\text{com}} = \sum_s \omega_{s,t} \boldsymbol{a}_{s,t}^{\text{com}}, \qquad \Delta \boldsymbol{f}_{b,t}^{\text{gran}} = \sum_s \omega_{s,t} \boldsymbol{a}_{s,t}^{\text{gran}}.$$

Using the price multiplier $\boldsymbol{\mathcal{M}}_t$ from Proposition 1, the corresponding price responses are $\Delta \boldsymbol{p}_{b,t}^k = \boldsymbol{\mathcal{M}}_t \Delta \boldsymbol{f}_{b,t}^k$ for $k \in \{\text{com}, \text{gran}\}$. Because the two components are orthogonal in holdings space but not necessarily after AUM weighting and price impact, we report a two-player Shapley split of the variance of $\Delta \boldsymbol{p}_{b,t}^{\text{com}} + \Delta \boldsymbol{p}_{b,t}^{\text{gran}}$.

Table 22. Common and granular components of top-tail behavioral tilts

	PCA of active tilts		Demand-system price variance			
Bias family	Sleeves	PC1 share	Common EW	Granular EW	Common VW	Granular VW
Overconfidence	10	51.1	21.8	78.2	30.4	69.6
		(1.1)	(3.4)	(3.4)	(3.7)	(3.7)
Lottery Preference	20	42.1	20.5	79.5	27.0	73.0
		(0.8)	(1.8)	(1.8)	(2.0)	(2.0)
Disposition Effect	20	34.6	9.9	90.1	12.0	88.0
		(0.7)	(2.7)	(2.7)	(3.0)	(3.0)
Extrapolation	20	51.9	11.9	88.1	12.8	87.2
		(1.5)	(4.2)	(4.2)	(4.3)	(4.3)

Notes: The table decomposes the top-tail household behavioral tilt within each bias family. For overconfidence, the top tail is the highest tercile in each wealth decile; for the other bias families, it is the top two quintiles in each wealth decile. Each sleeve's active tilt is its portfolio minus the AUM-weighted portfolio of all non-residual household sleeves in the same wealth decile. The PC1 column reports the time-series average share of equal-sleeve active-tilt variation explained by the first uncentered principal component. Common and granular columns AUM-weight the PC1 projection and residual across sleeves, push both flows through the demand-system inverse, and report the two-player Shapley share of the resulting equal-weighted or value-weighted price variance. Newey-West standard errors are in parentheses.

I.5 PCA of Top-Tail Bias Tilts

This subsection describes the diagnostic reported in Figure 5. For each bias family and month, we form the most-biased household sleeves within each wealth decile. For overconfidence, this gives one sleeve per wealth decile, the highest tercile, so there are 10 sleeves. For lottery preference, the disposition effect, and extrapolation, this gives two sleeves per wealth decile, the fourth and fifth quintiles, so there are 20 sleeves.

Let s index these sleeves and let $w(s)$ denote sleeve s's wealth decile. Let $\boldsymbol{w}_{s,t}$ be sleeve s's stock portfolio weights in month t, and let $\boldsymbol{w}_{w(s),t}^{\text{hh}}$ be the AUM-weighted portfolio of all

non-residual household sleeves in the same wealth decile. We define the sleeve's active tilt as

$$\boldsymbol{a}_{s,t} \equiv \boldsymbol{w}_{s,t} - \boldsymbol{w}^{\mathrm{hh}}_{w(s),t}.$$

The benchmark is therefore wealth-specific: each top-tail sleeve is compared to households with the same wealth rank.

For each bias family b and month t, stack these active tilts as columns of

$$\boldsymbol{A}^b_t \equiv \begin{bmatrix} \boldsymbol{a}_{1,t} & \cdots & \boldsymbol{a}_{S_b,t} \end{bmatrix},$$

where $S_b = 10$ for overconfidence and $S_b = 20$ for the other three bias families. We run an uncentered singular value decomposition,

$$\boldsymbol{A}^b_t = \boldsymbol{U}^b_t \boldsymbol{\Sigma}^b_t (\boldsymbol{V}^b_t)',$$

and define the cumulative K-component share as

$$C^b_{K,t} \equiv \frac{\sum_{j=1}^{K} (\sigma^b_{j,t})^2}{\sum_{j=1}^{S_b} (\sigma^b_{j,t})^2} = \frac{\|\boldsymbol{P}^b_{K,t} \boldsymbol{A}^b_t\|^2_F}{\|\boldsymbol{A}^b_t\|^2_F},$$

where $\boldsymbol{P}^b_{K,t}$ projects onto the first K left singular vectors. Figure 5 plots the time-series average of $C^b_{1,t}$ and the incremental second-component share $C^b_{2,t} - C^b_{1,t}$. The error bars are ± 2 Newey-West standard errors for the monthly series $C^b_{1,t}$ and $C^b_{2,t}$.

This object is descriptive. It is an equal-sleeve holdings diagnostic, not an AUM-weighted price-impact decomposition. A high value of $C^b_{K,t}$ means that a small number of stock-level tilt directions summarize the top-tail sleeve portfolios well; it does not by itself say how much those directions matter for equilibrium prices.

J Supporting Data and Auxiliary Diagnostics

This appendix collects supporting descriptives and auxiliary diagnostics that complement the main exhibits. It reports external anomaly returns, household-classification evidence,

bias-bin elasticity estimates, and validation tests for the household-level measures.

J.1 Value-Weighted Anomaly Returns

The main anomaly-attribution exercise emphasizes equal-weighted spreads because the behavioral mechanisms we study are cross-sectional predictions about stocks. Value-weighted spreads instead ask whether those mechanisms are strong among the largest firms. This is a useful conservative check, but weak value-weighted spreads do not imply that the underlying characteristic relation is absent from the cross-section. A value-weighted long–short portfolio can put substantial mass on a small number of large firms, so its return can reflect large-firm idiosyncratic variation as much as the broad anomaly mechanism. Recent evidence on aggregate market portfolios makes a related point: the standard value-weighted public-market factor contains a large public-firm component that is not compensated by broad macro risk (Byun et al., 2026).

Table 23 benchmarks this issue using the official U.S. factor returns from Jensen et al. (2023) for the 51 distinct anomalies that enter our creator predictions. We use the same February 2007–December 2016 window as the Norwegian expected-return analysis and average anomaly returns equally within each JKP family. Across all 51 anomalies, the equal-weighted portfolio earns 3.61% annually, compared with 1.76% for the capped-value-weighted portfolio; the corresponding Newey–West t-statistics are 1.99 and 1.23. The equal-weighted mean is also larger in six of the seven JKP families represented in the creator tests. These results support equal weighting as the more informative specification for the broad cross-sectional mechanisms studied in the paper.

Table 23. U.S. Returns for the JKP Anomalies in the Creator Tests

		Equal-weighted		Capped value-weighted	
JKP family	Anomalies	Mean (%)	t-stat	Mean (%)	t-stat
Intangibles	3	3.90	2.26	1.87	1.40
Investment	9	3.80	2.77	2.67	1.58
Momentum	6	2.75	0.84	0.33	0.11
New	6	3.15	0.85	1.49	0.52
Profitability	2	0.67	0.15	0.95	0.51
Trading frictions	12	2.84	0.86	1.74	0.57
Value	13	5.20	2.47	2.03	0.88
All creator anomalies	51	3.61	1.99	1.76	1.23

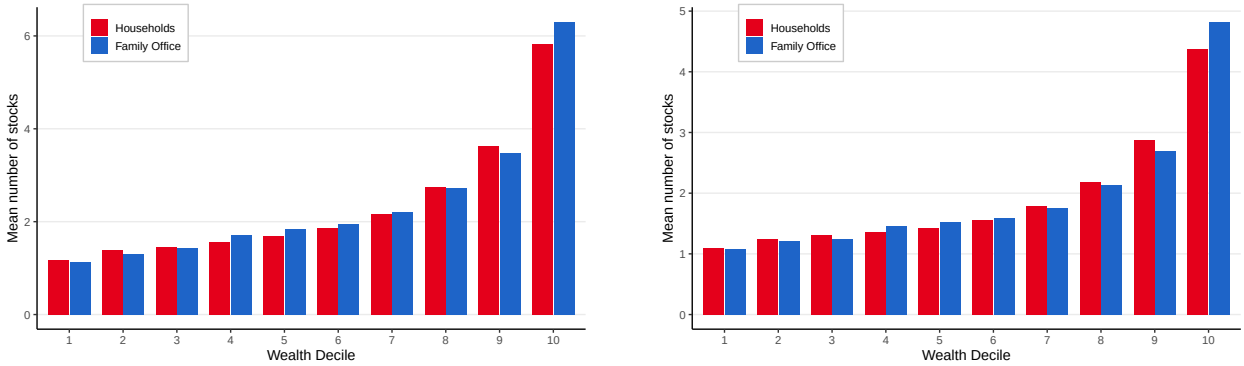
Notes: The table reports official U.S. monthly factor returns from the Global Factor Data library of Jensen et al. (2023) for the 51 distinct anomalies that enter the paper's creator predictions. The sample runs from February 2007 through December 2016. Within each month, returns are averaged equally across the anomalies in the indicated JKP family; the final row averages all 51 anomalies. Means are annualized and reported in percent. The t-statistics use Newey–West standard errors with three lags. The two columns use the library's equal-weighted and capped-value-weighted factor constructions, with each factor oriented according to its published return direction.

J.2 Household Classification Validation

We combine directly held household accounts with household-controlled private companies (family offices) in a single household sector. Figure 8 and Table 24 show that the two invest alike within wealth deciles. Within each decile, direct households and family offices hold a similar average number of stocks per account, and the cosine similarity between their average portfolio weights is high: it rises from 0.62 in the lowest wealth decile to about 0.90 in the top deciles, reaching 0.80 by the fourth decile. The two groups resemble each other most among wealthy households, whose holdings carry the most weight for prices, supporting their treatment as a single sector.

J.3 Bias-Bin Elasticity Estimates

The following tables report demand estimates separately across the bias bins used in the counterfactual exercises. They show the direct own-stock term, the factor-spillover block, and the implied stock-level elasticity summaries for each of the four bias families.



(a) Total Assets

(b) Inside Assets

Figure 8. Mean number of stocks by wealth decile, households vs. family offices. Time-series average of the cross-sectional mean number of stocks held per account, for direct households and for holdings traced through privately held companies (family offices), within each wealth decile (1 = lowest, 10 = highest). Panel (a) forms deciles on total assets; Panel (b) on inside (domestic) assets.

Table 24. Cosine Similarity: Direct Households vs. Family Offices

Wealth decile	Months	Mean	SD	Min	Max
D1	171	0.616	0.048	0.487	0.729
D2	171	0.729	0.037	0.632	0.798
D3	171	0.797	0.040	0.689	0.887
D4	171	0.800	0.029	0.716	0.860
D5	171	0.826	0.049	0.708	0.901
D6	171	0.849	0.029	0.764	0.900
D7	171	0.872	0.019	0.823	0.914
D8	171	0.897	0.014	0.859	0.936
D9	171	0.906	0.014	0.875	0.941
D10	171	0.892	0.020	0.843	0.936

Monthly cosine similarity between the average inside-asset portfolio weights of direct households and family offices within each wealth decile (1 = lowest, 10 = highest).

Panel A: Bias-bin coefficient summaries			
Statistic	T1	T2	T3
$\eta_{0,\mathcal{G}}$	-0.050	-0.091	0.081
	(0.058)	(0.057)	(0.050)
Γ_{11}	0.106	0.084	0.118
	(0.029)	(0.029)	(0.029)
Γ_{22}	0.028	-0.010	0.075
	(0.031)	(0.031)	(0.031)
Γ_{33}	0.001	0.006	0.050
	(0.029)	(0.030)	(0.031)
Γ_{44}	0.096	0.085	0.107
	(0.015)	(0.014)	(0.017)
Avg Off	0.002	0.006	0.000
	(0.010)	(0.010)	(0.010)
Min Off	-0.155	-0.148	-0.139
	(0.031)	(0.030)	(0.031)
Max Off	0.154	0.142	0.163
	(0.030)	(0.030)	(0.030)

Panel B: Implied stock-level elasticity averages			
Statistic	T1	T2	T3
Avg diag($\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$)	1.0523	1.0921	0.9223
	(0.0577)	(0.0566)	(0.0495)
Avg off-diag($\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$)	-0.000017	-0.000012	-0.000026
	(0.000005)	(0.000005)	(0.000005)

Table 25. Bias-bin elasticity summaries for Overconfidence. Each column reports the equal-weight average across household groups in the indicated reported bias bin. Panel A reports the direct own-stock term and spillover-matrix summaries. Panel B reports the implied stock-level average diagonal and average off-diagonal entries of $\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$. Standard errors are reported below the estimates. For the min/max rows, the standard error corresponds to the reported extremum within each household group before bin averaging.

Panel A: Bias-bin coefficient summaries					
Statistic	Q1	Q2	Q3	Q4	Q5
$\eta_{0,\mathcal{G}}$	-0.655	-0.322	0.029	0.338	0.414
	(0.082)	(0.075)	(0.060)	(0.043)	(0.039)
Γ_{11}	0.041	0.057	0.093	0.087	-0.081
	(0.014)	(0.018)	(0.020)	(0.024)	(0.027)
Γ_{22}	0.023	0.041	0.055	0.034	-0.022
	(0.016)	(0.017)	(0.016)	(0.016)	(0.022)
Γ_{33}	-0.058	-0.029	-0.004	0.045	0.123
	(0.018)	(0.023)	(0.028)	(0.030)	(0.030)
Γ_{44}	0.016	0.048	0.071	0.085	0.033
	(0.008)	(0.010)	(0.012)	(0.014)	(0.016)
Avg Off	-0.005	-0.004	0.007	0.017	-0.002
	(0.005)	(0.006)	(0.006)	(0.008)	(0.007)
Min Off	-0.081	-0.110	-0.071	-0.082	-0.144
	(0.014)	(0.019)	(0.020)	(0.020)	(0.027)
Max Off	0.083	0.094	0.098	0.140	0.112
	(0.021)	(0.020)	(0.022)	(0.027)	(0.025)

Panel B: Implied stock-level elasticity averages					
Statistic	Q1	Q2	Q3	Q4	Q5
Avg diag($\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$)	1.6554	1.3229	0.9728	0.6638	0.5868
	(0.0818)	(0.0745)	(0.0597)	(0.0425)	(0.0387)
Avg off-diag($\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$)	-0.000002	-0.000009	-0.000016	-0.000019	-0.000004
	(0.000002)	(0.000003)	(0.000003)	(0.000003)	(0.000004)

Table 26. Bias-bin elasticity summaries for Lottery Preference. Each column reports the equal-weight average across household groups in the indicated reported bias bin. Panel A reports the direct own-stock term and spillover-matrix summaries. Panel B reports the implied stock-level average diagonal and average off-diagonal entries of $\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$. Standard errors are reported below the estimates. For the min/max rows, the standard error corresponds to the reported extremum within each household group before bin averaging.

Panel A: Bias-bin coefficient summaries					
Statistic	Q1	Q2	Q3	Q4	Q5
$\eta_{0,\mathcal{G}}$	0.110	-0.243	-0.275	-0.194	0.330
	(0.051)	(0.067)	(0.065)	(0.060)	(0.045)
Γ_{11}	0.095	0.094	0.083	0.107	0.128
	(0.028)	(0.025)	(0.025)	(0.026)	(0.031)
Γ_{22}	0.068	0.021	0.003	0.020	0.144
	(0.029)	(0.024)	(0.023)	(0.026)	(0.031)
Γ_{33}	0.003	-0.046	-0.054	-0.047	0.117
	(0.032)	(0.027)	(0.027)	(0.029)	(0.033)
Γ_{44}	0.083	0.055	0.062	0.075	0.117
	(0.017)	(0.013)	(0.012)	(0.014)	(0.019)
Avg Off	0.019	0.018	0.011	0.013	-0.002
	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)
Min Off	-0.098	-0.093	-0.116	-0.124	-0.190
	(0.029)	(0.021)	(0.026)	(0.027)	(0.032)
Max Off	0.146	0.129	0.118	0.140	0.192
	(0.029)	(0.026)	(0.029)	(0.028)	(0.031)

Panel B: Implied stock-level elasticity averages					
Statistic	Q1	Q2	Q3	Q4	Q5
Avg diag($\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$)	0.8918	1.2437	1.2760	1.1950	0.6740
	(0.0511)	(0.0663)	(0.0651)	(0.0600)	(0.0451)
Avg off-diag($\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$)	-0.000018	-0.000009	-0.000007	-0.000011	-0.000038
	(0.000005)	(0.000004)	(0.000004)	(0.000004)	(0.000005)

Table 27. Bias-bin elasticity summaries for Disposition Effect. Each column reports the equal-weight average across household groups in the indicated reported bias bin. Panel A reports the direct own-stock term and spillover-matrix summaries. Panel B reports the implied stock-level average diagonal and average off-diagonal entries of $\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$. Standard errors are reported below the estimates. For the min/max rows, the standard error corresponds to the reported extremum within each household group before bin averaging.

Panel A: Bias-bin coefficient summaries					
Statistic	Q1	Q2	Q3	Q4	Q5
$\eta_{0,\mathcal{G}}$	0.483	0.495	0.434	0.491	0.601
	(0.043)	(0.044)	(0.047)	(0.049)	(0.049)
Γ_{11}	0.062	0.093	0.105	0.082	0.019
	(0.025)	(0.028)	(0.027)	(0.027)	(0.022)
Γ_{22}	0.085	0.132	0.124	0.073	0.040
	(0.026)	(0.023)	(0.023)	(0.025)	(0.022)
Γ_{33}	0.141	0.089	-0.026	0.005	0.007
	(0.032)	(0.032)	(0.031)	(0.032)	(0.030)
Γ_{44}	0.086	0.107	0.089	0.070	0.027
	(0.016)	(0.018)	(0.015)	(0.016)	(0.017)
Avg Off	0.002	-0.014	-0.016	-0.020	0.002
	(0.009)	(0.009)	(0.008)	(0.007)	(0.008)
Min Off	-0.135	-0.132	-0.150	-0.123	-0.120
	(0.028)	(0.030)	(0.030)	(0.029)	(0.026)
Max Off	0.125	0.113	0.120	0.109	0.122
	(0.027)	(0.027)	(0.028)	(0.025)	(0.029)

Panel B: Implied stock-level elasticity averages					
Statistic	Q1	Q2	Q3	Q4	Q5
Avg diag($\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$)	0.5198	0.5083	0.5682	0.5107	0.3995
	(0.0433)	(0.0433)	(0.0472)	(0.0491)	(0.0484)
Avg off-diag($\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$)	-0.000028	-0.000031	-0.000022	-0.000017	-0.000007
	(0.000004)	(0.000004)	(0.000004)	(0.000004)	(0.000004)

Table 28. Bias-bin elasticity summaries for extrapolation. Each column reports the equal-weight average across household groups in the indicated reported bias bin. Panel A reports the direct own-stock term and spillover-matrix summaries. Panel B reports the implied stock-level average diagonal and average off-diagonal entries of $\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t}$. Standard errors are reported below the estimates. For the min/max rows, the standard error corresponds to the reported extremum within each household group before bin averaging.

J.4 Bias-Measure Validation

A possible concern with how we measure behavioral tilts is that the sorts pick up something else. For example, households that load on the lottery, overconfidence, disposition, or extrapolation measures could potentially reflect investment skill or tax motivations. In what follows, we present several tests for the three measures whose household-level assignments are available. We measure the relevant gradient within a wealth decile and therefore hold wealth fixed.

J.4.1 Investment skill or bias?

One alternative explanation is that our measures capture investment skill rather than bias. Under this interpretation, the households we classify as most biased should outperform those we classify as least biased. For example, investors with high turnover or active share may be trading on private information. We test this hypothesis by forming a return spread between the most and least biased households. For each group, we use the actual stock holdings at month t and the subsequent stock returns from t to $t+1$. Within each wealth-decile-by-bias-bin cell, we value-weight household portfolio returns by assets under management. We then compute the return of the top bias bin minus that of the bottom bias bin within each wealth decile and month. Finally, we average these ten decile-specific spreads into a single monthly series and annualize the mean by multiplying by 12.

Table 29 reports the results. For the three measures that can be evaluated with the available household assignments, the most-biased minus least-biased return spread is negative: -2.34% per year for the disposition effect ($t = -2.6$), -6.54% for lottery preference, and -0.27% for overconfidence. The portfolio-level extrapolation archive does not retain the household returns needed for this test.

Table 29. Performance gradient: most- minus least-biased households, within wealth decile

Bias	Ann. active-return spread	NW t
Overconfidence	-0.27%	-0.53
Lottery Preference	-6.54%	-1.35
Disposition Effect	-2.34%	-2.58

Annualized difference in monthly portfolio return between the most- and least-biased bins, formed within each wealth decile and averaged across the ten deciles. Negative values mean that more-biased households earn lower returns than equally wealthy, less-biased households. The second column reports Newey–West t-statistics. Extrapolation is omitted because its available aggregate panels do not contain household-level returns or assignments.

J.4.2 Tax optimization or bias?

Overconfidence is measured by turnover, which raises the possibility that the high-turnover group sells for year-end tax reasons rather than because of overconfidence. To examine this possibility, we trace how the most overconfident households trade across the calendar year and compare the pattern against signatures of tax-loss selling. As in the main text, we sort households into overconfidence terciles each month and study the top tercile, whose turnover drives the measure. For every position held in the prior month, we record a sale when its share count falls, and we classify the stock as a winner or a loser by its price change over that month.

Our key assumption is that tax-loss harvesting shows up in the data via sales that concentrate in December and fall on losers. The reason is that only realized losses reduce the year's tax bill. In contrast, "true" overconfidence has no calendar pattern. We therefore report three statistics for the top tercile. First, we regress each household's monthly sell rate on a December indicator with year fixed effects. Tax-motivated selling implies a large positive December coefficient, while overconfidence implies a coefficient near zero. Second, we compute the loser share of sales in each calendar month—the fraction of sold positions whose price fell over the prior month; tax-loss harvesting implies a December peak in this share. Third, we measure the proportion of gains realized (PGR), the number of winning positions sold divided by the number of winning positions held, and the proportion of losses realized (PLR), the number of losing positions sold divided by the number of losing positions held. Tax-loss selling pushes PLR above PGR toward year-end, whereas overconfidence, which carries no calendar pattern, leaves PGR at or above PLR.

Table 30 reports the first measure—the December coefficient from the sell-rate regression, by tercile. Figure 9 presents the seasonality behind all three. Panel (a) plots the monthly sell rate for the three terciles, the series underlying the regression; panel (b) plots the loser share of sales, PGR, and PLR for the top tercile across calendar months.

Table 30. December selling by overconfidence tercile

Tercile	Non-Dec. mean	December coef.	(SE)	t
T1 (low)	0.0135	0.0011	(0.0001)	13.38
T2	0.0136	0.0011	(0.0001)	13.60
T3 (top)	0.0413	0.0020	(0.0001)	18.14

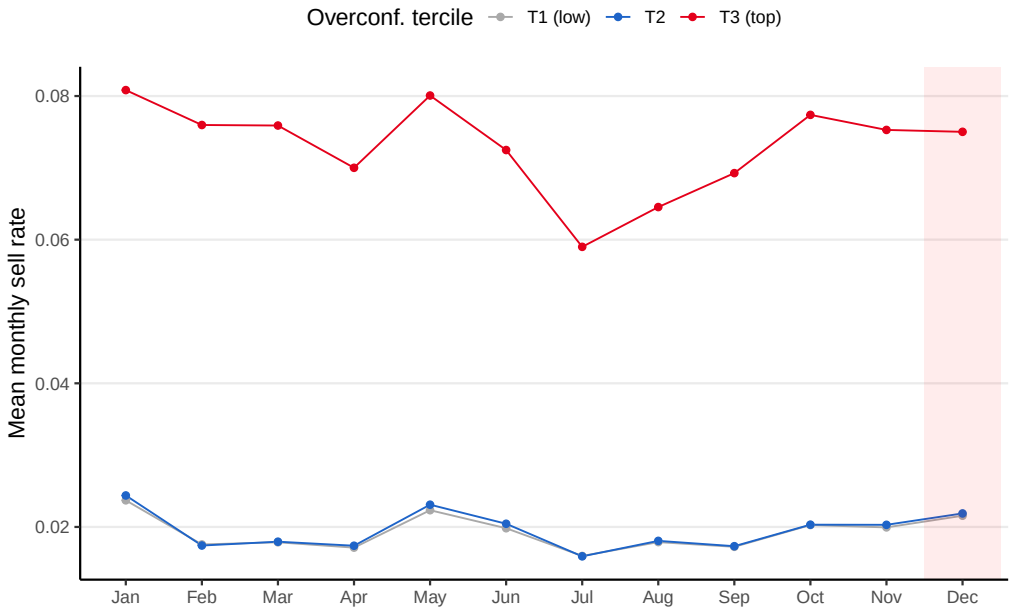
Household-month sell rate regressed on a December indicator with year fixed effects, standard errors clustered by household.

The table confirms that the sort isolates high-turnover households: the top tercile's baseline monthly sell rate, 0.041, is roughly three times that of the two lower terciles, 0.014. Yet its December coefficient is only 0.002, a rise of about 5% over its non-December rate, and, in proportional terms, no larger than the December increase in the low-turnover terciles (8%). The coefficient is statistically significant, but its magnitude is negligible. Tax-motivated selling would concentrate a December surge in exactly this high-turnover group, and we find none.

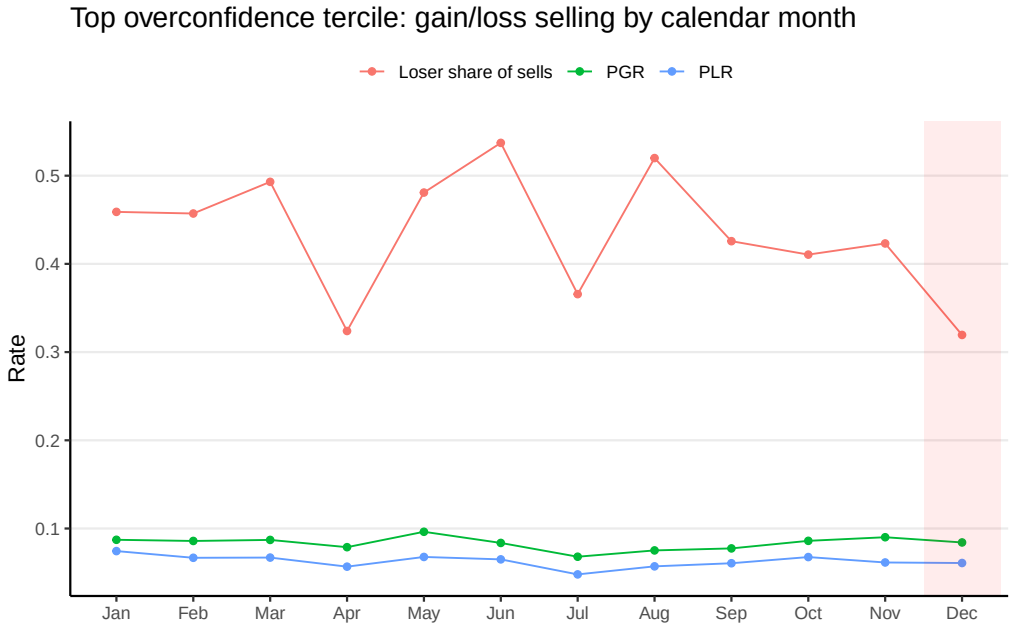
The figure sharpens the point. In panel (a), the top tercile sells more than the other terciles in every month, but its sell rate is flat across the calendar, with December sitting inside the normal monthly range rather than above it. Panel (b) shows the direction of those sales. For the top tercile, the loser share of sales is lowest in December, at 0.32, the lowest of any month, and the proportion of gains realized exceeds the proportion of losses realized in every month, including December. Year-end sales by the most overconfident households skew toward winners, not losers—the reverse of tax-loss harvesting. Overall, we conclude that the turnover measure is not due to tax timing.

J.4.3 Real constraints or bias?

A household holding essentially one stock, a founder or employer position, could score high on the volatility- and skewness-related measures without any behavioral tilt. A single undiversified position inherits the volatility and skewness of that one stock, so an investor



(a) Mean monthly sell rate by overconfidence tercile



(b) Gain/loss selling, top overconfidence tercile

Figure 9. December selling and the tax-timing placebo. December is shaded. Panel (a): mean household-month sell rate by calendar month for the three overconfidence terciles; the top tercile (T3) trades far more, but its December rate is not elevated relative to other months. Panel (b): for the top tercile, the loser share of sales, the proportion of gains realized (PGR), and the proportion of losses realized (PLR), by calendar month.

who fails to diversify can look like someone chasing lottery-like or high-variance payoffs. We test this by comparing portfolio concentration across bias bins within each wealth decile, using three measures: the largest weight on a particular stock in the household's portfolio (if the household has one stock, this would be 100%), the Herfindahl index of the household's portfolio weights, and the number of stocks it holds. If concentrated single positions, driven by a constraint unrelated to behavioral biases, drive the sorts, the most biased bins should be the most concentrated. Under the behavioral alternative, concentration is flat across bias bins, because the tilt reflects which stocks households hold rather than how many.

Table 31 matches the behavioral hypothesis for overconfidence, lottery preference, and the disposition effect. The weight in the single largest holding is flat or declining across their bias bins. For lottery preference, for example, it falls from 0.93 in the least-biased bin to 0.86 in the most-biased. The Herfindahl index and stock count tell the same story. Extrapolation is not included because the portfolio-level archive does not retain individual household positions.

Table 31. Portfolio concentration by bias bin

Bias	Bin 1 (low)	2	3	4	5 (high)
Panel A: Weight in Single Largest Holding					
Overconfidence	0.877	0.876	0.834		
Lottery Preference	0.934	0.862	0.820	0.817	0.860
Disposition Effect	0.839	0.891	0.889	0.850	0.824
Panel B: Herfindahl Index					
Overconfidence	0.844	0.843	0.795		
Lottery Preference	0.913	0.826	0.779	0.773	0.821
Disposition Effect	0.800	0.860	0.858	0.813	0.781
Panel C: Number of Stocks					
Overconfidence	1.85	1.86	2.36		
Lottery Preference	1.49	1.99	2.37	2.42	1.97
Disposition Effect	2.25	1.69	1.73	2.33	2.32

Portfolio concentration by bias bin, from inside (directly held) stock positions. Panel A reports the weight in each household's single largest holding, Panel B the Herfindahl index across its holdings, and Panel C the number of stocks held. Each cell is the time-series mean of the cross-sectional average within the bias bin. Overconfidence has three bins; lottery preference and the disposition effect have five. Extrapolation is omitted because its available aggregate panels do not retain individual household portfolios.

K Kernel-Similarity Robustness

The low-rank spillover block in (16) implies that the off-diagonal elasticity between stocks n and m is

$$\eta^{\text{tot}}_{\mathcal{G},t}(n,m) = \boldsymbol{x}'_{n,t}\Gamma_{\mathcal{G}}\boldsymbol{x}_{m,t}, \qquad n \neq m. \tag{54}$$

That structure is tractable, but it makes substitution monotone in characteristic levels rather than local in characteristic distance. For example, suppose we are considering bonds, and there is only one asset characteristic: the duration of the bond. To see the concern, suppose, for example, that $\Gamma_{\mathcal{G},kk} > 0$. Holding asset n fixed and suppose asset n is a 4-year duration bond. A 20-year bond can then look more substitutable than a 6-year bond simply because its duration score is larger, not because it is closer to a 4-year bond. In this appendix, we ask whether the three headline results survive when we replace that monotone spillover block with a similarity-based kernel specification.

For each month t, let $r_{n,k,t} \in [-0.5, 0.5]$ denote the raw rank-transformed value of characteristic $k \in \{\text{Book}, \text{Beta}, \text{ESG}, \text{ROE}\}$ for stock n. We construct the Gaussian similarity kernel

$$K_{k,t}(n,m) \equiv \phi\left(\frac{r_{n,k,t} - r_{m,k,t}}{\sigma_{k,t}}\right), \qquad K_{k,t}(n,n) = 0, \tag{55}$$

where $\sigma_{k,t}$ is the cross-sectional standard deviation of the ranked characteristic and $\phi(\cdot)$ is the standard normal density. Define the corresponding kernel price index by

$$y_{k,n,t} \equiv \sum_{m \neq n} K_{k,t}(n,m)\, g_{m,t}. \tag{56}$$

We then replace (19) and (20) with

$$w_{i,n,t} = \alpha_{i,t} + \sum_{k=1}^{4} r_{n,k,t}\beta_{k,i} + \eta^{\text{ker}}_{0,\mathcal{G}(i)} g_{n,t} + \sum_{k=1}^{4} \theta^{\text{ker}}_{\mathcal{G}(i),k} y_{k,n,t} + u_{i,n,t}, \tag{57}$$

$$R^{(12)}_{n,t\to t+12} = \alpha_n + \gamma_t + \sum_{k=1}^{4} r_{n,k,t}\delta_k + \beta^{(12),\text{ker}}_{R,0} g_{n,t} + \sum_{k=1}^{4} \vartheta^{(12),\text{ker}}_{R,k} y_{k,n,t} + \nu^{(12)}_{n,t}. \tag{58}$$

Thus the kernel elasticity and pass-through objects are

$$\boldsymbol{\eta}_{\mathcal{G},t}^{\text{tot,ker}} = \left(1 - \eta_{0,\mathcal{G}}^{\text{ker}}\right)\boldsymbol{I} - \sum_{k=1}^{4} \theta_{\mathcal{G},k}^{\text{ker}} \boldsymbol{K}_{k,t}, \qquad \boldsymbol{\Psi}_t^{\text{ker}} = \psi_0^{\text{ker}} \boldsymbol{I} + \sum_{k=1}^{4} \vartheta_k^{\text{ker}} \boldsymbol{K}_{k,t}, \tag{59}$$

and the expected-return pass-through matrix is again $\boldsymbol{\Theta}_t^{\text{ker}} = \boldsymbol{I} - \rho \boldsymbol{\Psi}_t^{\text{ker}}$. As in the main specification, we obtain its monthly coefficients from $\theta_0^{\text{ker}} = -\beta_{R,0}^{(12),\text{ker}}/12$ and $\vartheta_{\Theta,k}^{\text{ker}} = -\vartheta_{R,k}^{(12),\text{ker}}/12$, then recover the displayed $\boldsymbol{\Psi}_t^{\text{ker}}$. On the demand side, the own-price and kernel coefficients are still pooled within investor group exactly as in the main specification. The investor-level instrument for $g_{n,t}$ remains the leave-out-own Koijen–Yogo instrument, and each kernel price index is instrumented with the corresponding kernel-weighted instrument. Standard errors in Table 32 are clustered by stock and month. Table 33 reports the baseline rows together with the corresponding kernel rows, and Figure 10 compares the creator-prediction averages under the two systems.

Table 32 summarizes the resulting kernel demand and pass-through systems. The household own-price coefficient is -0.181, implying an average diagonal total elasticity of 1.181, while the average off-diagonal elasticity is 0.0619. On the pass-through side, the 12-month return specification delivers a monthly own pass-through coefficient of 0.008 (standard error 0.004), with an average off-diagonal component of 0.0001 in $\boldsymbol{\Theta}_t^{\text{ker}}$. Thus the kernel system keeps cross-stock substitution concentrated among nearby stocks in characteristic space and retains positive own-price pass-through.

Table 33 shows that the broad household variance result remains close to the main specification on the common February 2007–December 2016 sample. Under the R^2 Shapley decomposition, households account for 31.7% of equal-weighted cross-sectional return variance in the baseline and 31.4% in the kernel specification; in value-weighted terms the corresponding figures are 21.6% and 20.4%. Under the Standalone decomposition, the equal-weighted household share moves from 33.3% to 35.3%, while the value-weighted share moves from 19.1% to 22.8%. Within households, extrapolation's top-tail share is also stable, moving from 46.7% to 46.0% EW and from 48.0% to 48.2% VW.

Figure 10 shows positive creator-prediction averages for overconfidence, lottery preference, and the disposition effect under the panel-kernel system. Lottery preference remains

Panel A: Broad-group kernel demand system			
Group	$\beta_{0,\mathcal{G}}$	Avg diag($\eta_{\mathcal{G},t}$)	Avg off-diag($\eta_{\mathcal{G},t}$)
Households	-0.181	1.181	0.06188
	(0.177)	(0.177)	(0.02055)
Institutions	-0.042	1.042	0.04302
	(0.360)	(0.360)	(0.02503)
Governments	-1.526	2.526	0.12151
	(0.953)	(0.953)	(0.08736)
Listed firms	0.933	0.067	0.02480
	(0.309)	(0.309)	(0.03652)
Foreign	0.612	0.388	0.02632
	(0.171)	(0.171)	(0.01710)
Other	0.573	0.427	0.00419
	(0.251)	(0.251)	(0.03095)
Non-listed firms	0.650	0.350	0.02143
	(0.178)	(0.178)	(0.02201)

Panel B: Kernel pass-through system		
Matrix	Own	Avg off-diag
Ψ_t	0.995	-0.0001
	(0.004)	(0.0004)
Θ_t	0.008	0.0001
	(0.004)	(0.0004)

Table 32. Panel-kernel demand and pass-through summaries. Panel A reports the broad-group own-price coefficient and the implied average diagonal and average off-diagonal entries of the elasticity matrix. Panel B reports the corresponding own and average off-diagonal terms for the pass-through matrix and the implied one-step expected-return pass-through matrix. Standard errors are reported below the estimates.

Panel A: Household variance share				
Specification	R^2 Shapley EW	R^2 Shapley VW	Standalone EW	Standalone VW
Baseline	31.7	21.6	33.3	19.1
	(2.7)	(2.3)	(4.1)	(2.9)
Kernel	31.4	20.4	35.3	22.8
	(1.6)	(1.3)	(1.8)	(1.4)

Panel B: Concentration in the most-biased sleeves				
Bias family / spec	R^2 Shapley EW	R^2 Shapley VW	Standalone EW	Standalone VW
Overconfidence: Baseline	69.9	65.8	75.9	82.2
	(4.4)	(2.9)	(4.1)	(0.9)
Overconfidence: Kernel	73.6	70.6	79.6	82.4
	(2.1)	(2.2)	(2.3)	(2.7)
Lottery Preference: Baseline	89.8	70.8	86.4	69.6
	(2.9)	(5.0)	(3.2)	(0.8)
Lottery Preference: Kernel	88.9	67.3	87.3	69.2
	(0.5)	(1.3)	(0.5)	(1.1)
Disposition Effect: Baseline	61.2	49.5	62.3	54.7
	(7.2)	(4.7)	(7.2)	(0.6)
Disposition Effect: Kernel	61.0	46.6	61.8	48.8
	(1.2)	(2.2)	(1.6)	(2.9)
Extrapolation: Baseline	46.7	48.0	46.9	43.2
	(6.8)	(5.4)	(7.0)	(0.6)
Extrapolation: Kernel	46.0	48.2	46.4	49.7
	(4.3)	(3.5)	(4.5)	(3.9)

Table 33. Variance-share robustness under the panel-kernel specification. Panel A reports the household share of cross-sectional return variance in the broad decomposition. Panel B reports the same concentration object as Table 9: the top tercile for overconfidence and the top two quintiles for the other bias families. Each panel reports the baseline and panel-kernel estimates on the common February 2007–December 2016 sample for the R^2 Shapley and Standalone decompositions. Standard errors are reported below the estimates.

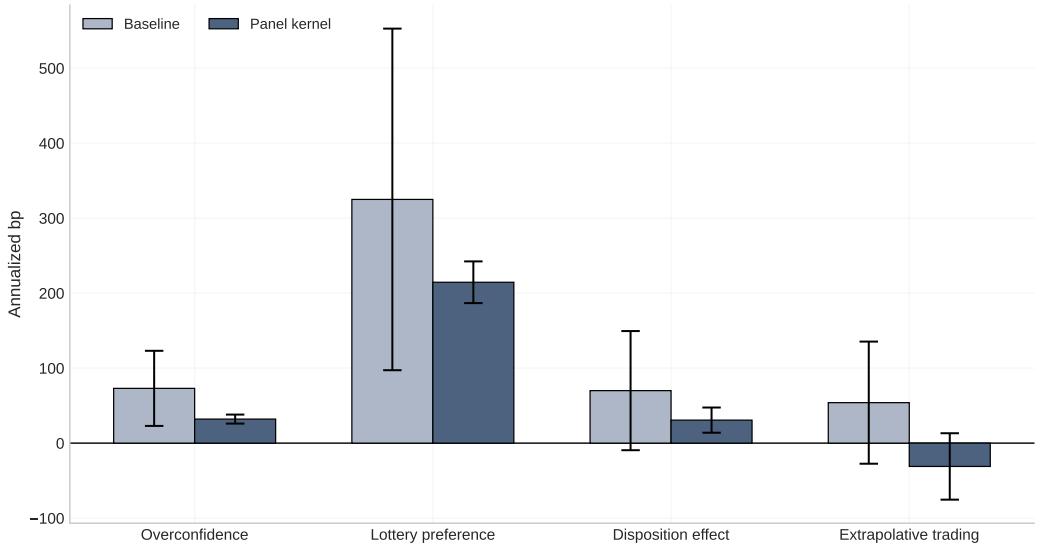


Figure 10. Average creator contribution, baseline versus panel-kernel robustness. Each bar averages the equal-weighted contribution across the creator predictions for that behavioral tilt, reported on the common 20% AUM scale. The panel-kernel specification makes cross-stock substitution depend on local similarity in characteristic space. Error bars report 95% confidence intervals, using 5,000 common joint draws for the baseline and Newey–West standard errors from the monthly panel-kernel series.

the largest contribution. The extrapolation estimate is -31.1 annualized basis points.

L Cross-Sectional IV Robustness

The Koijen–Yogo instrument was originally designed for cross-sectional IV estimation, while we use it in the main analysis within a panel-IV demand system. A natural concern is therefore that the identifying restriction may be easier to defend when demand estimation is carried out month by month. At the same time, the low-rank factor spillover specification from Section 3.4 cannot be estimated cross section by cross section, because within a given month the interacted spillover regressors collapse into linear combinations of the stock characteristics as discussed by Haddad et al. (2025). We therefore re-estimate demand month by month under the kernel-similarity specification from Appendix K, and then average the monthly coefficients over time in the spirit of Fama–MacBeth. Pass-through is inherently a resolution-horizon object, so a separate 12-month coefficient cannot be identified from a single cross section. We pair the monthly demand estimates with the pooled 12-month panel-kernel pass-through estimate from Appendix K; this exercise consequently isolates the effect of estimating demand in monthly cross sections.

Fix either the broad investor-group system or one of the bias-specific systems, and let $\mathcal{G}_s$ denote the groups in that system. In the bias-specific cross-sectional systems, we pool households within the reported bias bins and keep the nonhousehold groups unchanged, so the monthly systems line up with the concentration and reallocation objects that we report in the main text. For each month t, we estimate

$$w_{i,n,t} = \alpha_{i,t} + \sum_{k=1}^{4} r_{n,k,t}\beta_{k,i} + \sum_{g\in\mathcal{G}_s} 1\{i \in g\}\beta^{\text{cs}}_{0,g,t} g_{n,t} + \sum_{k=1}^{4} \theta^{\text{cs}}_{k,t} y_{k,n,t} + u_{i,n,t}, \tag{60}$$

where $r_{n,k,t} \in [-0.5, 0.5]$ is the ranked stock characteristic from (55), $y_{k,n,t}$ is the kernel price index from (56), and the common kernel block $(\theta^{\text{cs}}_{1,t}, \ldots, \theta^{\text{cs}}_{4,t})$ is pooled within the monthly system. The group-specific own-price terms are still identified from the interacted leave-out-own Koijen–Yogo instrument, and each kernel price index is instrumented with the corresponding kernel-weighted instrument.

Let $\hat{v}^{\theta}_{k,t}$ denote the estimated variance of the monthly kernel demand coefficient. As in Appendix I.3, we leave the group-specific own-price terms unshrunk and apply empirical

Bayes only to the four demand spillover coefficients:

$$\tilde{\theta}^{\text{cs}}_{k,t} \equiv \frac{\left(\tau^{\text{cs}}_{\eta}\right)^2}{\left(\tau^{\text{cs}}_{\eta}\right)^2 + \hat{v}^{\theta}_{k,t}} \hat{\theta}^{\text{cs}}_{k,t}. \tag{61}$$

We choose $\left(\tau^{\text{cs}}_{\eta}\right)^2$ jointly across the broad demand system and the four bias-specific demand systems by marginal likelihood, exactly as in the main specification. We then average the monthly demand coefficients over time:

$$\bar{\beta}^{\text{cs}}_{0,g} \equiv \frac{1}{T_g} \sum_t \hat{\beta}^{\text{cs}}_{0,g,t}, \qquad \bar{\tilde{\theta}}^{\text{cs}}_{k} \equiv \frac{1}{T} \sum_t \tilde{\theta}^{\text{cs}}_{k,t}. \tag{62}$$

The resulting elasticity and pass-through objects are

$$\boldsymbol{\eta}^{\text{tot,cs}}_{g,t} = \left(1 - \bar{\beta}^{\text{cs}}_{0,g}\right)\boldsymbol{I} - \sum_{k=1}^{4} \bar{\tilde{\theta}}^{\text{cs}}_{k} \boldsymbol{K}_{k,t}, \qquad \boldsymbol{\Psi}^{\text{cs}}_{t} = \boldsymbol{\Psi}^{\text{ker}}_{t}, \qquad \boldsymbol{\Theta}^{\text{cs}}_{t} = \boldsymbol{\Theta}^{\text{ker}}_{t}, \tag{63}$$

So this robustness keeps the local-similarity spillover structure from Appendix K, imposes the demand-side IV restriction only within monthly cross sections, and uses the same empirical-Bayes logic as the main specification to discount noisy monthly demand estimates before forming the time-averaged system. It holds the pooled kernel pass-through system fixed.

Table 34 summarizes the resulting cross-sectional-kernel demand coefficients and the pooled kernel pass-through used with them. The household own-price coefficient is −0.008, implying an average diagonal total elasticity of 1.008. After month-level empirical-Bayes shrinkage, the average off-diagonal entry of the broad-group elasticity matrix is 0.00675. Panel B repeats the pooled 12-month panel-kernel pass-through estimate, with a monthly own coefficient of 0.008, because the cross-sectional exercise varies demand estimation rather than the return-resolution horizon.

Table 35 compares the common-sample baseline with the cross-sectional-kernel analogue. Households account for 33.3% of equal-weighted variance and 23.6% of value-weighted variance under the R^2 Shapley decomposition, compared with 31.7% and 21.6% in the baseline. Under the Standalone decomposition, the corresponding shares are 33.2% and 20.7%. For extrapolation, the top-tail R^2 Shapley shares are 43.7% EW and 46.7% VW, close to the

Panel A: Broad-group cross-sectional-kernel demand system			
Group	$\beta_{0,\mathcal{G}}$	Avg diag($\eta_{\mathcal{G},t}$)	Avg off-diag($\eta_{\mathcal{G},t}$)
Households	-0.008	1.008	0.00675
	(0.192)	(0.192)	(0.00173)
Institutions	0.394	0.606	0.00675
	(0.302)	(0.302)	(0.00173)
Governments	-1.396	2.396	0.00675
	(0.211)	(0.211)	(0.00173)
Listed firms	0.941	0.059	0.00675
	(0.206)	(0.206)	(0.00173)
Foreign	0.822	0.178	0.00675
	(0.222)	(0.222)	(0.00173)
Other	0.652	0.348	0.00675
	(0.200)	(0.200)	(0.00173)
Non-listed firms	0.754	0.246	0.00675
	(0.195)	(0.195)	(0.00173)

Panel B: Cross-sectional-kernel pass-through system		
Matrix	Own	Avg off-diag
Ψ_t	0.995	-0.00005
	(0.004)	(0.00037)
Θ_t	0.008	0.00005
	(0.004)	(0.00037)

Table 34. Cross-sectional-kernel demand and pass-through summaries. Each month, the demand equation is estimated in the cross section; its kernel spillover coefficients are shrunk with month-level empirical Bayes and then averaged over time, while the own-price terms are left unshrunk. Because resolving returns require a time-series horizon, this demand robustness uses the pooled panel-kernel pass-through estimate from Appendix K, based on realized compounded 12-month returns and expressed in monthly units. Panel A reports the broad-group own-price coefficient and the implied elasticity summaries. Panel B reports the pooled kernel price-law and expected-return pass-through matrices. Demand standard errors are Newey–West; pass-through standard errors are two-way clustered by stock and month.

Panel A: Household variance share				
Specification	R^2 Shapley EW	R^2 Shapley VW	Standalone EW	Standalone VW
Baseline	31.7	21.6	33.3	19.1
	(2.7)	(2.3)	(4.1)	(2.9)
CS kernel	33.3	23.6	33.2	20.7
	(8.3)	(4.7)	(11.0)	(8.9)

Panel B: Concentration in the most-biased sleeves				
Bias family / spec	R^2 Shapley EW	R^2 Shapley VW	Standalone EW	Standalone VW
Overconfidence: Baseline	69.9	65.8	75.9	82.2
	(4.4)	(2.9)	(4.1)	(0.9)
Overconfidence: CS kernel	47.0	42.5	60.8	61.5
	(10.9)	(10.2)	(26.5)	(26.6)
Lottery Preference: Baseline	89.8	70.8	86.4	69.6
	(2.9)	(5.0)	(3.2)	(0.8)
Lottery Preference: CS kernel	46.4	49.6	82.6	69.4
	(11.2)	(11.6)	(16.7)	(19.1)
Disposition Effect: Baseline	61.2	49.5	62.3	54.7
	(7.2)	(4.7)	(7.2)	(0.6)
Disposition Effect: CS kernel	62.4	60.8	59.0	54.4
	(16.2)	(15.4)	(25.5)	(24.3)
Extrapolation: Baseline	46.7	48.0	46.9	43.2
	(6.8)	(5.4)	(7.0)	(0.6)
Extrapolation: CS kernel	43.7	46.7	41.0	47.4
	(12.4)	(12.0)	(21.0)	(19.7)

Table 35. Variance-share robustness under the cross-sectional-kernel demand specification. Each month, demand is estimated in the cross section, the monthly demand spillover coefficients are shrunk with empirical Bayes, and the resulting demand system is then averaged over time. Panel A reports the household share of cross-sectional return variance in the broad decomposition. Panel B reports the same concentration object as Table 9: the top tercile for overconfidence and the top two quintiles for the other bias families. Each panel reports the baseline and cross-sectional-kernel estimates on the common February 2007–December 2016 sample for the R^2 Shapley and Standalone decompositions. Standard errors are reported below the estimates.

main estimates.

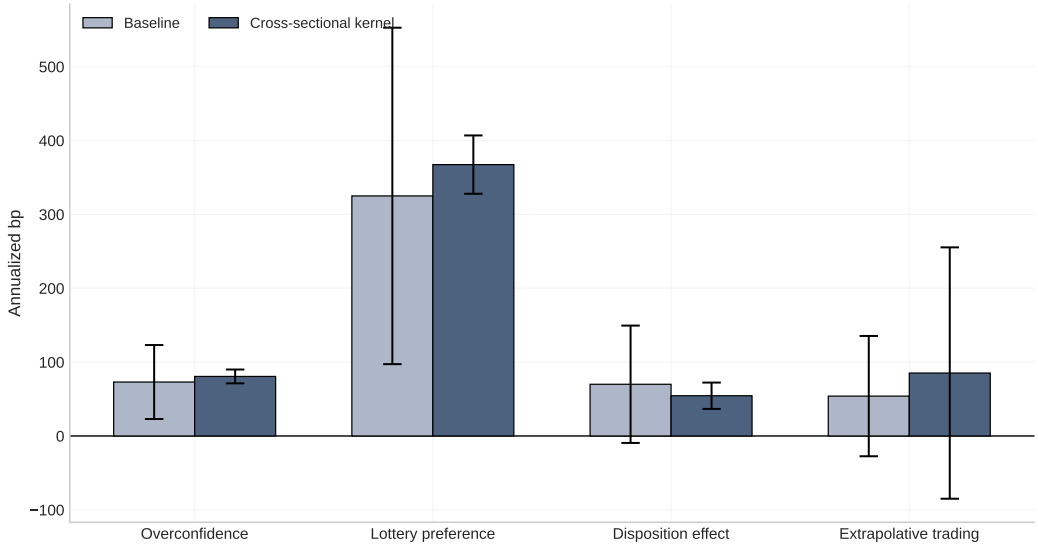


Figure 11. Average creator contribution, baseline versus cross-sectional-kernel robustness. Each bar averages the equal-weighted contribution across the creator predictions for that behavioral tilt, reported on the common 20% AUM scale. The cross-sectional-kernel system estimates demand month by month and combines the time-averaged demand coefficients with the panel-kernel pass-through estimate. Error bars report 95% confidence intervals, using 5,000 common joint draws for the baseline and Newey–West standard errors from the monthly cross-sectional-kernel series.

Figure 11 shows positive creator-prediction averages for all four behavioral tilts. The estimates are 80.6, 367.4, 54.5, and 85.2 annualized basis points for overconfidence, lottery preference, the disposition effect, and extrapolation, respectively.

M Elasticity and Pass-Through Diagnostics

As Davis et al. (2026) emphasizes, the total elasticity is made up of sensitivity to expected returns and risk effects. In the main analysis, we use the total elasticity, but it is also reasonable that the relevant counterfactual could involve different pass-through dynamics or endogenous risk effects perceived by investors. In this appendix, we show that the main results are robust to variation in those objects. We first recover the structural objects, namely the sensitivities to expected returns and risk effects. We then use those objects in counterfactuals to show that the main results are robust to variation in the pass-through

and risk responses. For a longer discussion of this type of exercise, see Davis et al. (2026).

M.1 Recovery Setup

Start from our structural demand system in (46). For investor i, this is

$$\boldsymbol{q}_{i,t} = a_{i,t}\mathbf{1} + \bar{\boldsymbol{q}}_{i,t} + \boldsymbol{C}_{i,t}\boldsymbol{\mu}_t - \sum_{k=1}^{4} \boldsymbol{D}_{i,k,t}\boldsymbol{\beta}_{k,t}v_{k,t} - \boldsymbol{E}_{i,t}\boldsymbol{\sigma}_t^2 + \boldsymbol{\xi}_{i,t}. \tag{64}$$

Combining this with the expected-return and risk mappings in (11) and (47) gives the total-elasticity object

$$\boldsymbol{\eta}_{i,t}^{\text{tot}} = \boldsymbol{C}_{i,t}\boldsymbol{\Theta}_t + \sum_{k=1}^{4} \boldsymbol{D}_{i,k,t}\boldsymbol{\beta}_{k,t}\boldsymbol{\gamma}_{k,t}^{V\prime} + \boldsymbol{E}_{i,t}\boldsymbol{\Gamma}_t^{\sigma}. \tag{65}$$

The recovery keeps this notation but does not attempt to identify the raw stock-by-stock objects $\boldsymbol{C}_{i,t}$, $\boldsymbol{D}_{i,k,t}$, and $\boldsymbol{E}_{i,t}$ separately. Instead, it decomposes the expected-return, factor-risk, and idiosyncratic-risk channels in (65) into a grouped low-rank component plus latent terms that allow subjective beliefs about returns and risk to have richer stock-level structure. Here N_t is the number of stocks in month t, $\boldsymbol{X}_t \in \mathbb{R}^{N_t \times 4}$ is the orthonormal characteristic basis from Section 3.4, $\mathcal{G}(i)$ denotes investor i's broad group, and

$$\boldsymbol{\Theta}^F \equiv \theta_0 \boldsymbol{I}_4 + \Gamma_\Theta \tag{66}$$

is the factor-space pass-through matrix associated with $\boldsymbol{\Theta}_t$. We also impose the same low-rank factor representation used above, $\boldsymbol{\beta}_t = \boldsymbol{X}_t$, together with the restricted risk-pass-through laws

$$\boldsymbol{\Gamma}_t^V = G^V \boldsymbol{X}_t', \qquad \boldsymbol{\Gamma}_t^{\sigma} = \gamma_\sigma \boldsymbol{I}_{N_t}. \tag{67}$$

For each broad group $\mathcal{G}$, we write these three channels as

$$\boldsymbol{C}_{\mathcal{G},t}\boldsymbol{\Theta}_t = a_{C,\mathcal{G}}\boldsymbol{\Theta}_t + \boldsymbol{X}_t\boldsymbol{H}_{\mathcal{G}}\boldsymbol{\Theta}^F\boldsymbol{X}_t' + \widetilde{\boldsymbol{U}}_{\mathcal{G},t}^{C}, \tag{68}$$

$$\sum_{k=1}^{4} \boldsymbol{D}_{\mathcal{G},k,t}\boldsymbol{\beta}_{k,t}v_{k,t} = \boldsymbol{X}_t\boldsymbol{D}_{\mathcal{G}}G^V\boldsymbol{X}_t' + \widetilde{\boldsymbol{U}}_{\mathcal{G},t}^{D}, \tag{69}$$

$$\boldsymbol{E}_{\mathcal{G},t}\boldsymbol{\sigma}_t^2 = e_{\mathcal{G}}\gamma_\sigma \boldsymbol{I}_{N_t} + \widetilde{\boldsymbol{U}}_{\mathcal{G},t}^{E}, \tag{70}$$

where $\boldsymbol{H}_{\mathcal{G}}$ is symmetric, $\boldsymbol{D}_{\mathcal{G}} \equiv \text{diag}(d_{1,\mathcal{G}}, \ldots, d_{4,\mathcal{G}})$, and the latent terms $\widetilde{\boldsymbol{U}}^{C}_{\mathcal{G},t}$, $\widetilde{\boldsymbol{U}}^{D}_{\mathcal{G},t}$, and $\widetilde{\boldsymbol{U}}^{E}_{\mathcal{G},t}$ capture higher-dimensional expected-return and risk beliefs that are not summarized by the grouped low-rank component. Let

$$\widetilde{\boldsymbol{U}}_{\mathcal{G},t} \equiv \widetilde{\boldsymbol{U}}^{C}_{\mathcal{G},t} - \widetilde{\boldsymbol{U}}^{D}_{\mathcal{G},t} - \widetilde{\boldsymbol{U}}^{E}_{\mathcal{G},t}. \tag{71}$$

We then decompose this combined latent term into its low-rank projection and orthogonal remainder,

$$\widetilde{\boldsymbol{U}}_{\mathcal{G},t} = u_{0,\mathcal{G}} \boldsymbol{I}_{N_t} + \boldsymbol{X}_t \boldsymbol{U}_{\mathcal{G}} \boldsymbol{X}_t' + \boldsymbol{U}^{\perp}_{\mathcal{G},t}, \tag{72}$$

with orthogonality conditions

$$\text{tr}\big(\boldsymbol{U}^{\perp}_{\mathcal{G},t}\big) = 0, \qquad \boldsymbol{X}_t' \boldsymbol{U}^{\perp}_{\mathcal{G},t} \boldsymbol{X}_t = \boldsymbol{0}. \tag{73}$$

So the next subsection estimates the risk mappings and then recovers the low-rank coefficients $(a_{C,\mathcal{G}}, \boldsymbol{H}_{\mathcal{G}}, \boldsymbol{D}_{\mathcal{G}}, e_{\mathcal{G}}, u_{0,\mathcal{G}}, \boldsymbol{U}_{\mathcal{G}})$ group by group.

M.2 Recovery of Expected-Return and Risk Components

We first estimate the risk mappings in (67). For each factor-variance coordinate $m = 1, \ldots, 4$, let $mvar_{m,t+1}$ denote the realized variance of factor return m over month $t+1$, constructed as the sum of squared daily IPCA factor returns over that month. We then run the monthly IV projection

$$mvar_{m,t+1} = \alpha_m + \sum_{b=1}^{4} g^{V}_{mb} z_{b,t} + u_{m,t+1}, \tag{74}$$

where α_m is an equation-specific intercept, $\boldsymbol{z}_t \equiv \boldsymbol{X}_t' \boldsymbol{g}_t$ collects the month-t factor-portfolio valuation gaps, with $\boldsymbol{g}_t = \boldsymbol{p}_t - \boldsymbol{b}_t$ as defined earlier, and $\hat{\boldsymbol{z}}_t \equiv \boldsymbol{X}_t' \hat{\boldsymbol{g}}_t$ is the IV fitted counterpart constructed from the instrumented valuation-gap vector $\hat{\boldsymbol{g}}_t$ used in the pass-through estimation. Stacking the four equations yields $\hat{G}^{V}$. Second, we use the following return-based

idiosyncratic-risk regression to get $\hat{\gamma}_\sigma$:

$$\log RV^{\text{idio}}_{n,t+1} = \alpha^\sigma_n + \delta^\sigma_t + \gamma_\sigma r_{n,t} + u^\sigma_{n,t+1}, \tag{75}$$

where $RV^{\text{idio}}_{n,t+1}$ is stock n's realized idiosyncratic variance over month $t+1$, $r_{n,t}$ is its month-t return, α^σ_n is a stock fixed effect, and δ^σ_t is a month fixed effect. This gives $\hat{\gamma}_\sigma \approx -0.17$. Table 36 reports the factor-risk first stage.

Response	PC1 gap	PC2 gap	PC3 gap	PC4 gap
PC1	-0.0143	-0.0162	-0.0171	-0.0160
	(0.1398)	(0.1070)	(0.0617)	(0.0890)
PC2	0.0009	-0.0029	-0.0131	0.0130
	(0.0813)	(0.1730)	(0.1006)	(0.0779)
PC3	-0.0061	-0.0089	-0.0051	-0.0059
	(0.0360)	(0.1119)	(0.0609)	(0.0425)
PC4	-0.0055	-0.0046	-0.0002	-0.0024
	(0.0330)	(0.0194)	(0.0128)	(0.0125)

Table 36. Next-month factor-variance pass-through estimates used in the structural recovery. Each row reports the IV projection of a next-month factor-variance coordinate on the current factor-valuation gaps. The standard errors come from the 12-month moving-block bootstrap used in the recovery exercise.

We then combine these risk estimates with the baseline pass-through matrix from Table 7. Let

$$\hat{\boldsymbol{\Theta}}^F \equiv \hat{\theta}_0 \boldsymbol{I}_4 + \hat{\Gamma}_\Theta \tag{76}$$

denote the estimated factor-space pass-through matrix from the main specification in Table 7, and let

$$\hat{\boldsymbol{\eta}}^F_{\mathcal{G}} \equiv (1 - \hat{\eta}_{0,\mathcal{G}}) \boldsymbol{I}_4 + \hat{\Gamma}^\eta_{\mathcal{G}} \tag{77}$$

denote the factor-space total elasticity implied by the broad-group estimate in Table 6. Here $\hat{\theta}_0$ is the estimated own-stock pass-through coefficient, $\hat{\Gamma}_\Theta$ is the estimated 4×4 factor-spillover matrix from Table 7, $\hat{\eta}_{0,\mathcal{G}}$ is the group's estimated own-stock total-elasticity coefficient, and $\hat{\Gamma}^\eta_{\mathcal{G}}$ is the group's estimated factor-spillover matrix from Table 6. Define the own total-elasticity coefficient

$$b_{0,\mathcal{G}} \equiv 1 - \hat{\eta}_{0,\mathcal{G}}, \tag{78}$$

and write the unknown structural coefficient vector as

$$\phi_{\mathcal{G}} \equiv \begin{bmatrix} a_{C,\mathcal{G}} \\ h_{\mathcal{G}} \\ d_{\mathcal{G}} \\ e_{\mathcal{G}} \end{bmatrix}, \tag{79}$$

where $h_{\mathcal{G}}$ stacks the 10 unique entries of the symmetric matrix $\boldsymbol{H}_{\mathcal{G}}$ and $d_{\mathcal{G}} = (d_{1,\mathcal{G}}, \ldots, d_{4,\mathcal{G}})'$. Substituting (68)–(72) into (65) and using the estimated mappings gives

$$\boldsymbol{\eta}^{\text{tot}}_{\mathcal{G},t} = a_{C,\mathcal{G}}\boldsymbol{\Theta}_t + \boldsymbol{X}_t\boldsymbol{H}_{\mathcal{G}}\hat{\boldsymbol{\Theta}}^{F}\boldsymbol{X}'_t - \boldsymbol{X}_t\boldsymbol{D}_{\mathcal{G}}\hat{G}^{V}\boldsymbol{X}'_t - e_{\mathcal{G}}\hat{\gamma}_{\sigma}\boldsymbol{I}_{N_t} + u_{0,\mathcal{G}}\boldsymbol{I}_{N_t} + \boldsymbol{X}_t\boldsymbol{U}_{\mathcal{G}}\boldsymbol{X}'_t + \boldsymbol{U}^{\perp}_{\mathcal{G},t}, \tag{80}$$

This is the stock-level total elasticity counterpart for equation (80). If we take the trace of (80) we get:

$$b_{0,\mathcal{G}} = \hat{\theta}_0 a_{C,\mathcal{G}} - \hat{\gamma}_{\sigma} e_{\mathcal{G}} + u_{0,\mathcal{G}}, \tag{81}$$

If we pre-multiply (80) by $\boldsymbol{X}'_t$ and post-multiply by $\boldsymbol{X}_t$, we get the factor-space projection:

$$\hat{\Gamma}^{\eta}_{\mathcal{G}} = a_{C,\mathcal{G}}\hat{\Gamma}_{\Theta} + \boldsymbol{H}_{\mathcal{G}}\hat{\boldsymbol{\Theta}}^{F} - \boldsymbol{D}_{\mathcal{G}}\hat{G}^{V} + \boldsymbol{U}_{\mathcal{G}}, \tag{82}$$

Note that $\boldsymbol{U}^{\perp}_{\mathcal{G},t}$ drops out of both equations given our assumptions above. These two equations serve as the key equations we use to identify the expected return and risk components sensitivity parameters in demand.

Let M_H map $h_{\mathcal{G}}$ into $\text{vec}(\boldsymbol{H}_{\mathcal{G}})$ and let M_D map $d_{\mathcal{G}}$ into $\text{vec}(\boldsymbol{D}_{\mathcal{G}})$. Vectorizing (82) gives

$$\text{vec}(\hat{\Gamma}^{\eta}_{\mathcal{G}}) = \text{vec}(\hat{\Gamma}_{\Theta})\, a_{C,\mathcal{G}} + \big((\hat{\boldsymbol{\Theta}}^{F})' \otimes \boldsymbol{I}_4\big) M_H h_{\mathcal{G}} - \big((\hat{G}^{V})' \otimes \boldsymbol{I}_4\big) M_D d_{\mathcal{G}} + \text{vec}(\boldsymbol{U}_{\mathcal{G}}). \tag{83}$$

Stack the total-elasticity target and residual as

$$y_{\mathcal{G}} \equiv \begin{bmatrix} b_{0,\mathcal{G}} \\ \text{vec}(\hat{\Gamma}^{\eta}_{\mathcal{G}}) \end{bmatrix}, \qquad u_{\mathcal{G}} \equiv \begin{bmatrix} u_{0,\mathcal{G}} \\ \text{vec}(\boldsymbol{U}_{\mathcal{G}}) \end{bmatrix}, \tag{84}$$

so the stacked recovery system is

$$y_{\mathcal{G}} = M\phi_{\mathcal{G}} + u_{\mathcal{G}}, \tag{85}$$

with design matrix

$$M \equiv \begin{bmatrix} \hat{\theta}_0 & 0 & 0 & -\hat{\gamma}_\sigma \\ \mathrm{vec}(\hat{\Gamma}_\Theta) & ((\boldsymbol{\Theta}^F)' \otimes \boldsymbol{I}_4)M_H & -((\hat{G}^V)' \otimes \boldsymbol{I}_4)M_D & 0 \end{bmatrix}. \tag{86}$$

We interpret this recovery as a projection of the estimated total elasticity onto the restricted structural class spanned by $(a_{C,\mathcal{G}}, \boldsymbol{H}_{\mathcal{G}}, \boldsymbol{D}_{\mathcal{G}}, e_{\mathcal{G}})$, with the leftover absorbed by $(u_{0,\mathcal{G}}, \boldsymbol{U}_{\mathcal{G}})$. We then estimate $\phi_{\mathcal{G}}$ by ridge minimum distance,

$$\hat{\phi}_{\mathcal{G}} \in \arg\min_{\phi} \left\{ \begin{array}{l} (y_{\mathcal{G}} - M\phi)' W_{\mathcal{G}} (y_{\mathcal{G}} - M\phi) \\ + \lambda_a q_{a,\mathcal{G}} \left(a_{C,\mathcal{G}} - \dfrac{b_{0,\mathcal{G}}}{\hat{\theta}_0} \right)^2 \\ + \lambda_H h'_{\mathcal{G}} Q_{H,\mathcal{G}} h_{\mathcal{G}} + \lambda_H d'_{\mathcal{G}} Q_{D,\mathcal{G}} d_{\mathcal{G}} \\ + \lambda_e q_{e,\mathcal{G}} \left(e_{\mathcal{G}} - \rho_e \dfrac{b_{0,\mathcal{G}}}{-\hat{\gamma}_\sigma} \right)^2 \end{array} \right\}, \tag{87}$$

where $W_{\mathcal{G}}$ is the inverse covariance matrix of the empirical-Bayes total-elasticity coefficients for group $\mathcal{G}$. Let $\mathrm{diag}(M'W_{\mathcal{G}}M)$ be partitioned into the corresponding scalar and block entries $q_{a,\mathcal{G}}$, $Q_{H,\mathcal{G}}$, $Q_{D,\mathcal{G}}$, and $q_{e,\mathcal{G}}$. So the recovery shrinks the flexible expected-return and factor-risk blocks toward zero, shrinks the own expected-return coefficient toward the no-risk benchmark $b_{0,\mathcal{G}}/\hat{\theta}_0$, and shrinks the idiosyncratic-risk coefficient toward a positive fraction ρ_e of $b_{0,\mathcal{G}}/(-\hat{\gamma}_\sigma)$. The ridge penalty parameters $(\lambda_H, \lambda_a, \lambda_e, \rho_e)$ are chosen by minimizing the average generalized cross-validation criterion (GCV score)[22]. The selected values are

[22]For each broad group $\mathcal{G}$, the recovery computes

$$\mathrm{GCV}_{\mathcal{G}} = \frac{\big((y_{\mathcal{G}} - M\hat{\phi}_{\mathcal{G}})' W_{\mathcal{G}} (y_{\mathcal{G}} - M\hat{\phi}_{\mathcal{G}})\big)/n_{\mathcal{G}}}{\big(1 - \mathrm{tr}(S_{\mathcal{G}})/n_{\mathcal{G}}\big)^2},$$

where $n_{\mathcal{G}} = \dim(y_{\mathcal{G}})$ and

$$S_{\mathcal{G}} = M\big(M'W_{\mathcal{G}}M + \Lambda_{\mathcal{G}}\big)^{-1} M'W_{\mathcal{G}}$$

is the ridge hat matrix, with $\Lambda_{\mathcal{G}}$ the diagonal penalty matrix implied by $(\lambda_H, \lambda_a, \lambda_e)$. We evaluate this score over a grid of candidate values, average it across the seven broad groups, and then choose the lowest average among the candidates that also satisfy $e_{\mathcal{G}} > 0$ and a positive broad-group average stock-level diagonal of $\boldsymbol{C}_{\mathcal{G},t}$ for every group.

$\hat{\lambda}_H \approx 39.81$, $\hat{\lambda}_a = 10^7$, $\hat{\lambda}_e = 10^8$, and $\hat{\rho}_e = 0.10$.

To compute the standard errors reported in Tables 37 and 38, we redraw the broad-group total-elasticity coefficients from their estimated covariance matrices, conditional on the pass-through and risk mappings. For Table 39, we instead use the same bootstrap as in the main text: each draw uses the corresponding bootstrapped demand and pass-through estimates, holds the recovery tuning parameters fixed, and then applies the scenario map. Figure 12 reports Newey–West standard errors from the monthly creator-prediction averages under each recovered system.

M.3 Recovered Structural Objects

Table 37 reports the recovered expected-return block. Panel A gives the factor-space coefficients, while Panel B maps the recovered object back into stock space. Under the baseline recovery, the average stock-level diagonal of $\boldsymbol{C}_{\mathcal{G},t}$ is positive for every broad group. For households, that average diagonal is 37.8, the average signed off-diagonal is -0.0019, and the average absolute off-diagonal is 0.251. The same pattern, a positive diagonal with much smaller signed cross-stock terms, holds for the other broad groups.

Table 38 reports the recovered risk block. The factor-risk component remains close to zero on average in stock space for every broad investor group, while the idiosyncratic-risk term stays positive for every group. For households, the recovered own idiosyncratic-risk term is 0.5155, compared with an average own factor-risk term of 0.0007. The other broad groups show the same qualitative pattern: small average factor-risk contributions in stock space and positive own idiosyncratic-risk terms. Thus the positivity-feasible recovery assigns the diagonal of total elasticity to a combination of a positive expected-return block and a positive idiosyncratic-risk block, not to large average factor-risk elasticities.

M.4 Counterfactual Pass-Through and Endogenous-Risk Robustness

The recovered structural blocks are useful because they let us vary the dynamic and risk mappings without re-estimating the demand system. This exercise differs from the main

Panel A: Factor-Space Expected-Return Coefficients					
Group	a_C	h_{11}	h_{22}	h_{33}	h_{44}
Households	37.548	9.272	21.717	-7.315	1.736
	(5.480)	(9.947)	(19.247)	(5.335)	(2.209)
Institutions	36.426	-4.722	-2.844	-20.080	3.552
	(12.998)	(2.481)	(1.664)	(4.909)	(1.099)
Governments	83.519	-2.625	-24.557	-9.108	-1.566
	(29.695)	(1.231)	(6.825)	(3.357)	(0.523)
Listed firms	4.392	-0.217	1.366	-0.209	-0.329
	(12.469)	(0.503)	(0.721)	(1.625)	(0.342)
Foreign	12.269	-0.462	1.484	3.683	2.067
	(5.670)	(0.875)	(0.476)	(1.473)	(0.616)
Other	15.262	-1.634	0.208	0.010	0.474
	(8.556)	(1.095)	(0.554)	(1.803)	(0.290)
Non-listed firms	11.529	-0.144	-1.223	-0.736	0.383
	(6.224)	(0.762)	(1.659)	(0.662)	(0.942)

Panel B: Stock-Level $\boldsymbol{C}_{\mathcal{G},t}$ Summaries			
Group	Avg diag($\boldsymbol{C}_{\mathcal{G},t}$)	Avg off-diag($\boldsymbol{C}_{\mathcal{G},t}$) ($\times$1000)	Avg $\lvert$off-diag$\rvert$ ($\times$1000)
Households	37.766	-1.8829	251.051
	(5.591)	(1.9244)	(199.066)
Institutions	36.219	1.7854	158.212
	(12.939)	(0.5972)	(26.921)
Governments	83.194	2.8051	213.131
	(29.605)	(0.8085)	(40.652)
Listed firms	4.397	-0.0453	22.779
	(12.459)	(0.1265)	(8.717)
Foreign	12.327	-0.5018	107.387
	(5.652)	(0.1837)	(9.560)
Other	15.254	0.0698	26.885
	(8.536)	(0.1964)	(12.965)
Non-listed firms	11.515	0.1275	16.435
	(6.213)	(0.1574)	(5.350)

Table 37. Recovered expected-return demand loadings by broad investor group. Panel A reports the scalar own-expected-return coefficient $a_{C,\mathcal{G}}$ and the diagonal entries of the symmetric factor-space matrix $\boldsymbol{H}_{\mathcal{G}}$ in $\boldsymbol{C}_{\mathcal{G},t} = a_{C,\mathcal{G}}\boldsymbol{I} + \boldsymbol{X}_t\boldsymbol{H}_{\mathcal{G}}\boldsymbol{X}_t'$. The recovery projects the estimated total elasticity onto this restricted structural class, allows a residual, shrinks $\boldsymbol{H}_{\mathcal{G}}$ and the factor-risk block toward zero, shrinks $a_{C,\mathcal{G}}$ toward $b_{0,\mathcal{G}}/\hat{\theta}_0$, and shrinks $e_{\mathcal{G}}$ toward a positive fraction of $b_{0,\mathcal{G}}/(-\hat{\gamma}_\sigma)$. Panel B maps the recovered object into stock space and reports the average diagonal, signed off-diagonal, and absolute off-diagonal entry of $\boldsymbol{C}_{\mathcal{G},t}$ across stocks and months. Standard errors redraw only the broad-group total-elasticity coefficients while holding the pass-through matrix, the risk mappings, and the recovery tuning parameters fixed.

Panel A: Recovered Risk Coefficients						
Group	d_1	d_2	d_3	d_4	e	$-e\gamma_\sigma$
Households	-0.073	-2.708	11.239	15.130	2.776	0.516
	(2.052)	(1.413)	(15.791)	(3.632)	(0.403)	(0.075)
Institutions	-0.126	-1.889	29.728	-3.352	1.249	0.232
	(0.740)	(0.789)	(9.469)	(3.622)	(0.565)	(0.105)
Governments	-1.146	-15.848	8.445	3.559	6.183	1.148
	(0.457)	(5.838)	(4.234)	(1.534)	(2.200)	(0.409)
Listed firms	0.333	0.645	0.999	-0.089	0.325	0.060
	(0.420)	(1.630)	(1.479)	(0.695)	(0.925)	(0.172)
Foreign	1.089	7.121	-32.074	9.350	0.906	0.168
	(0.704)	(1.222)	(6.759)	(1.592)	(0.421)	(0.078)
Other	0.704	-0.874	0.464	1.011	1.133	0.210
	(0.578)	(1.240)	(2.848)	(0.938)	(0.635)	(0.118)
Non-listed firms	0.533	0.320	-6.123	0.156	0.854	0.159
	(0.692)	(1.120)	(3.168)	(1.841)	(0.462)	(0.086)

Panel B: Stock-Level Risk Summaries					
Group	Own factor (×1000)	Cross factor (×1000)	|Cross factor| (×1000)	Own idio	Own total
Households	0.728	-0.006	1.192	0.5155	0.5163
	(0.505)	(0.004)	(1.030)	(0.0749)	(0.0751)
Institutions	1.155	-0.010	2.466	0.2319	0.2331
	(0.346)	(0.003)	(0.546)	(0.1049)	(0.1049)
Governments	-0.098	0.001	1.954	1.1482	1.1481
	(0.139)	(0.001)	(0.698)	(0.4086)	(0.4086)
Listed firms	0.099	-0.001	0.130	0.0604	0.0605
	(0.086)	(0.001)	(0.147)	(0.1717)	(0.1717)
Foreign	-0.882	0.008	2.864	0.1682	0.1673
	(0.311)	(0.003)	(0.543)	(0.0782)	(0.0780)
Other	0.106	-0.001	0.189	0.2104	0.2105
	(0.143)	(0.001)	(0.209)	(0.1179)	(0.1179)
Non-listed firms	-0.189	0.002	0.517	0.1586	0.1584
	(0.167)	(0.001)	(0.173)	(0.0858)	(0.0858)

Table 38. Recovered risk loadings by broad investor group. Panel A reports the diagonal factor-risk coefficients $d_{1,\mathcal{G}}, \ldots, d_{4,\mathcal{G}}$, the scalar idiosyncratic-risk coefficient $e_{\mathcal{G}}$, and the implied own idiosyncratic-risk term $-e_{\mathcal{G}}\gamma_\sigma$. Panel B reports the corresponding stock-level summaries of the recovered risk component $-\boldsymbol{X}_t\boldsymbol{D}_{\mathcal{G}}\hat{G}^V\boldsymbol{X}_t' - e_{\mathcal{G}}\gamma_\sigma\boldsymbol{I}$, again averaged across all stocks and months. The first three Panel B columns are shown in (×1000) units, while the own-idiosyncratic and own-total columns remain in original units. Standard errors redraw only the broad-group total-elasticity coefficients from their estimated covariance matrices while holding the pass-through matrix, the risk mappings, and the recovery tuning parameters fixed.

expected-return attribution, which holds the total elasticity fixed. Here we instead recover $(\hat{\boldsymbol{C}}_{g,t}, \hat{\boldsymbol{D}}_g, \hat{e}_g)$ group by group under our main pass-through matrix, keep the recovery residual, and then let the total elasticity move mechanically when pass-through or endogenous risk changes. Here g denotes the relevant recovery group: a broad investor group in the broad-group exercise and a bias bin in the bias-specific exercises. For the broad-group result, we do this for the seven broad investor groups. For the bias-bin variance and anomaly exercises, we apply the same recovery separately within each bias system and then hold those recovered objects fixed across scenarios.

Formally, for scenario s we set

$$\boldsymbol{\Theta}_t^{(s)} = \lambda_s \hat{\boldsymbol{\Theta}}_t, \qquad G^{V,(s)} = \kappa_s \hat{G}^V, \qquad \gamma_\sigma^{(s)} = \kappa_s \hat{\gamma}_\sigma, \tag{88}$$

where $\hat{\boldsymbol{\Theta}}_t$ is the estimated stock-space pass-through matrix from Table 7, λ_s scales pass-through, and κ_s scales both endogenous-risk mappings. We then define the implied total elasticity as

$$\boldsymbol{\eta}_{g,t}^{\text{tot},(s)} = \hat{\boldsymbol{C}}_{g,t}\boldsymbol{\Theta}_t^{(s)} - \boldsymbol{X}_t \hat{\boldsymbol{D}}_g G^{V,(s)} \boldsymbol{X}_t' - \hat{e}_g \gamma_\sigma^{(s)} \boldsymbol{I} + \hat{u}_{0,g}\boldsymbol{I} + \boldsymbol{X}_t \hat{\boldsymbol{U}}_g \boldsymbol{X}_t', \tag{89}$$

Equivalently, if $\hat{\boldsymbol{\Theta}}^F \equiv \hat{\theta}_0 \boldsymbol{I}_4 + \hat{\Gamma}_\Theta$, $\hat{\boldsymbol{C}}_{g,t} = \hat{a}_{C,g}\boldsymbol{I} + \boldsymbol{X}_t \hat{\boldsymbol{H}}_g \boldsymbol{X}_t'$, and $\hat{u}_g$ denotes the stacked residual from the ridge recovery, then the scenario-specific total-elasticity coefficients are

$$y_g^{(s)} = M^{(s)}\hat{\phi}_g + \hat{u}_g = y_g + (M^{(s)} - M)\hat{\phi}_g. \tag{90}$$

Thus the robustness exercise simply varies pass-through or endogenous risk, maps those changes into a new total-elasticity matrix, and then recomputes the counterfactual objects. When $(\lambda_s, \kappa_s) = (1, 1)$, this construction returns the estimated total-elasticity matrix for every stock and month.

We consider five scenarios: the recovery baseline $(\lambda_s, \kappa_s) = (1, 1)$; a low-pass-through case with $\lambda_s = 0.008/\hat{\theta}_0$; a high-pass-through case with $\lambda_s = 0.048/\hat{\theta}_0$; and low- and high-risk cases with $\kappa_s = 0.8$ and $\kappa_s = 1.2$. The pass-through targets are taken directly from the literature discussed in Section 4.1: the low target of 0.008 is the monthly counterpart

of the annual estimate in Koijen et al. (2022), the high target of 0.048 is from Davis et al. (2025b), and Davis (2025) report an intermediate monthly estimate of about 0.01. For each scenario, we recompute the three headline counterfactual objects with the implied $\boldsymbol{\eta}_{g,t}^{\mathrm{tot},(s)}$: the broad variance-share decomposition, the within-bias variance concentration exercise, and the within-bias reallocation anomaly counterfactual. The baseline row uses the common February 2007–December 2016 sample and the estimated total-elasticity matrix without rescaling either pass-through or risk.

Table 39 shows that the first two headline results are stable across the recovery scenarios. The household share of modeled variance ranges from 20.8–35.5% EW and 16.7–22.9% VW. The value-weighted extrapolation top-tail share ranges only from 47.6–49.5%. Thus its contribution to the household variance channel does not depend on the chosen pass-through or risk mapping.

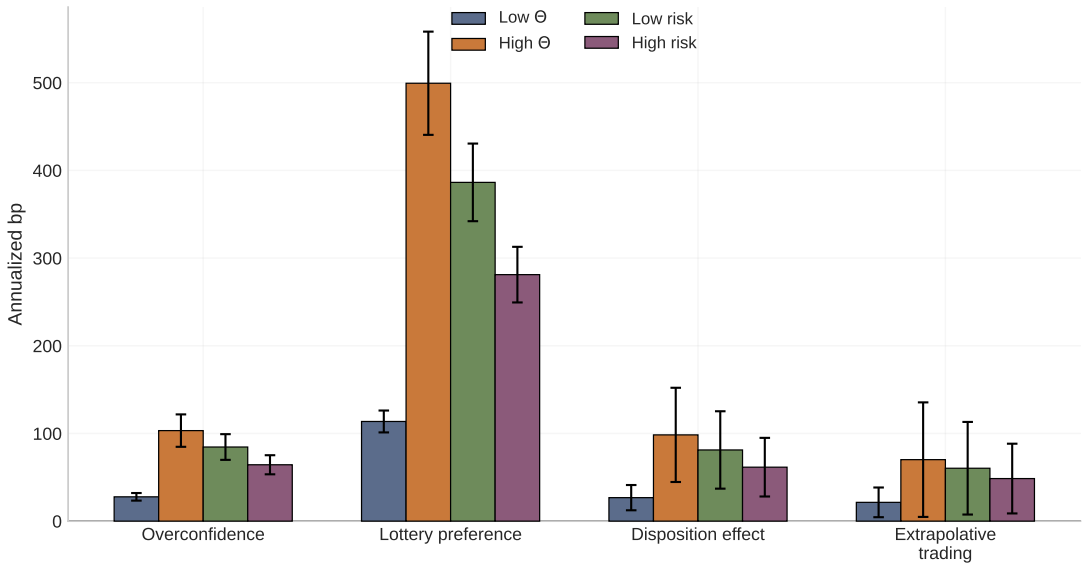


Figure 12. Recovery-based robustness of average creator contributions. Each bar averages the equal-weighted contribution across the creator predictions for that behavioral tilt, reported on the common 20% AUM scale. The four bars correspond to low and high pass-through and low and high endogenous risk. In every case we hold the recovered structural demand blocks fixed and vary only the pass-through and risk mappings. Error bars report 95% confidence intervals based on Newey–West standard errors from the monthly creator-prediction averages.

Figure 12 shows positive creator-prediction averages for all four behavioral tilts in every structural counterfactual. Lottery preference remains the largest contribution throughout.

Panel A: Household Variance Share		
Scenario	HH Owen EW	HH Owen VW
Baseline	31.7	21.6
	(2.6)	(2.1)
Low Θ (0.008)	20.8	16.7
	(2.8)	(2.3)
High Θ (0.048)	35.5	22.9
	(3.0)	(2.4)
Low risk (0.8×)	30.7	21.1
	(2.6)	(2.1)
High risk (1.2×)	32.6	22.1
	(2.7)	(2.2)

Panel B: Top-Tail Within-Bias Variance Share (VW)				
Scenario	Overconfidence	Lottery	Disposition	Extrap. trading
Baseline	65.8	70.8	49.5	48.0
	(2.9)	(4.9)	(4.4)	(5.5)
Low Θ (0.008)	66.9	69.5	46.9	49.5
	(3.4)	(5.4)	(4.5)	(5.5)
High Θ (0.048)	65.0	71.5	50.3	47.6
	(5.3)	(5.9)	(4.4)	(5.5)
Low risk (0.8×)	65.5	70.7	49.5	47.9
	(3.0)	(5.0)	(4.4)	(5.5)
High risk (1.2×)	66.1	70.9	49.5	48.1
	(3.0)	(4.8)	(4.4)	(5.5)

Table 39. Recovery-based robustness of the first two headline results on the common February 2007–December 2016 sample. In every row we hold the recovered structural coefficients and their recovery residual fixed and recompute total elasticities under alternative pass-through and endogenous-risk mappings. The low- and high-pass-through rows scale the baseline pass-through matrix so that its own term equals 0.008 and 0.048, respectively. The low- and high-risk rows scale both $\hat{G}^V$ and $\hat{\gamma}_\sigma$ by 0.8 and 1.2. At the baseline scenario, the recovered object coincides with the estimated total elasticity for every stock and month. Panel A reports the household share of modeled variance under the hierarchical-Owen decomposition. Panel B reports the value-weighted share of within-bias household variance coming from the most-biased sleeve, defined as $T3$ for overconfidence and $Q4 + Q5$ for the three quintile-based sorts. Standard errors come from the same joint bootstrap used in the main text, with the recovery tuning parameters held fixed within each draw.

The extrapolation estimate ranges from 21.4 to 70.1 annualized basis points as pass-through and endogenous risk vary.

N Dynamic Portfolio-Maintenance Robustness

The anomaly counterfactuals are one-step objects: after the month-t within-bias reallocation, we calculate the month-t portfolios' next-month removal effects and report their negatives—the observed-minus-counterfactual contributions. A natural follow-up question is how the creator-prediction averages evolve if the initial distortion persists and the anomaly portfolios are rebalanced. We study that dynamic implementation question using the same 50% reallocation shock as in (40).

Fix a bias family b and an origin month t. The initial flow and price distortion are exactly those in (40) and (41). We then hold the month-t empirical-Bayes elasticity and pass-through objects fixed and propagate the resulting distortion forward using the price law of motion in (10). Writing the month-t persistence matrix as $\hat{\boldsymbol{\Psi}}_t$ and the implied pass-through matrix as $\hat{\boldsymbol{\Theta}}_t = \boldsymbol{I} - \rho\hat{\boldsymbol{\Psi}}_t$, the horizon-$s$ price distortion and removal-effect vector are

$$\Delta^{\text{cf}}\boldsymbol{p}^b_{t+s-1|t} \equiv \hat{\boldsymbol{\Psi}}_t^{\,s-1}\Delta^{\text{cf}}\boldsymbol{p}^b_t, \tag{91}$$

$$\Delta^{\text{cf}}\boldsymbol{\mu}^b_{t+s|t} \equiv -\hat{\boldsymbol{\Theta}}_t\hat{\boldsymbol{\Psi}}_t^{\,s-1}\Delta^{\text{cf}}\boldsymbol{p}^b_t, \qquad s = 1,\dots,24. \tag{92}$$

As in the main analysis, the reported stock-level contribution is the negative of the removal effect,

$$\boldsymbol{\mathcal{C}}^b_{t+s|t} \equiv -\Delta^{\text{cf}}\boldsymbol{\mu}^b_{t+s|t} = \hat{\boldsymbol{\Theta}}_t\hat{\boldsymbol{\Psi}}_t^{\,s-1}\Delta^{\text{cf}}\boldsymbol{p}^b_t. \tag{93}$$

Thus the dynamic exercise stays within the local linearization, but asks what the initial month-t shock implies for observed-minus-counterfactual expected-return contributions up to two years ahead.

We report two portfolio mappings of this same expected-return path. The blue line in Figure 13 keeps each formation-month anomaly portfolio fixed:

$$\mathcal{C}^{P,\text{fixed}}_{b,t+s|t} \equiv 12\,(\boldsymbol{W}^P_t)'\boldsymbol{\mathcal{C}}^b_{t+s|t}. \tag{94}$$

The red line instead reconstitutes each anomaly portfolio monthly using its standard month-specific assignment:

$$\mathcal{C}^{P,\text{reb}}_{b,t+s|t} \equiv 12\,(\boldsymbol{W}^{P}_{t+s-1})'\boldsymbol{\mathcal{C}}^{b}_{t+s|t}. \tag{95}$$

Here $\boldsymbol{W}^{P}_{t+s-1}$ is the equal-weight long–short portfolio formed in month $t+s-1$ using the same 30/70 breakpoint rule and published orientation as in the main analysis, restricted to stocks in the formation-month price system. This is a dynamic portfolio-maintenance exercise, not a second equilibrium in which the maintenance trades themselves feed back through a new price-impact system. At each horizon, we first average equally across the creator predictions within a behavioral tilt.

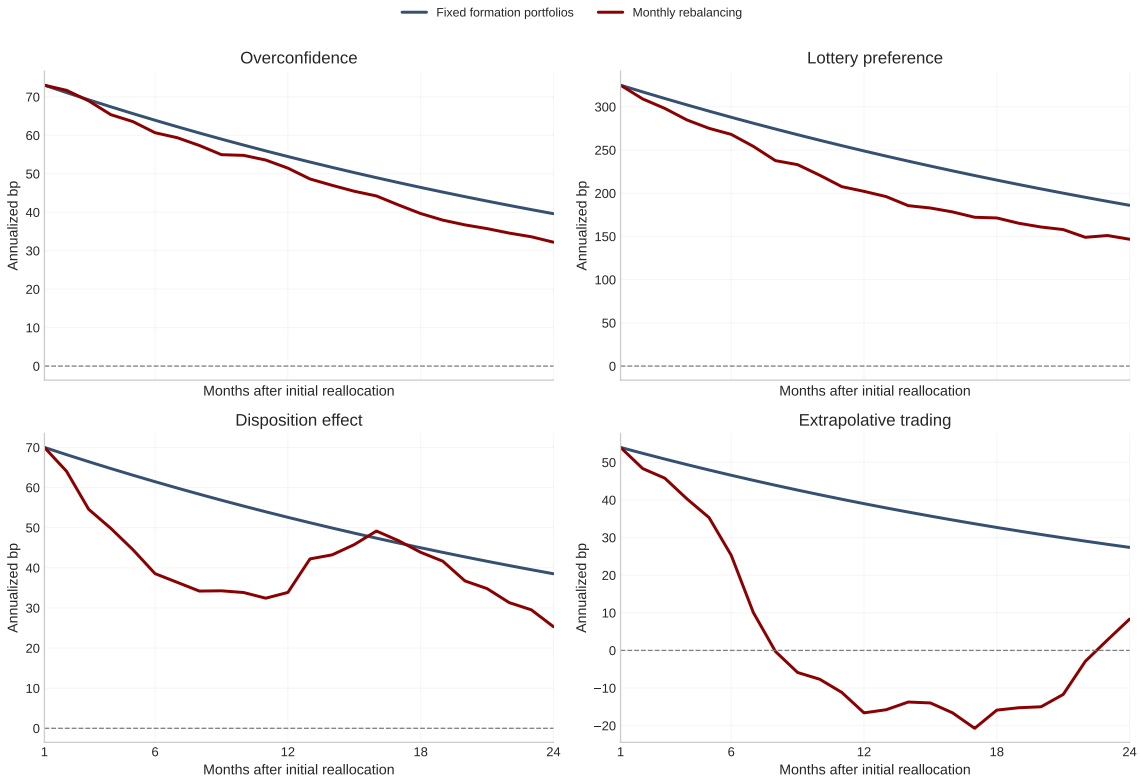


Figure 13. Dynamic robustness of average creator contributions. Each panel averages equally across the creator predictions for that behavioral tilt and reports contributions on the common 20% AUM scale. The paths start from the same 50% within-bias reallocation used in the main analysis. The blue line keeps each formation-month anomaly portfolio fixed after the initial shock. The red line reconstitutes each anomaly portfolio monthly using its standard month-specific assignment.

Figure 13 shows that the two paths coincide at horizon 1 by construction. The fixed-

formation creator average remains positive for every behavioral tilt throughout the 24-month horizon, with lottery preference the largest contribution. Monthly rebalancing preserves positive averages throughout for overconfidence, lottery preference, and the disposition effect. For extrapolation, the average starts at 54.0 annualized basis points; the rebalanced path moves around zero at intermediate horizons, and the month-24 estimates are 27.4 basis points for fixed formation portfolios and 8.3 basis points with monthly rebalancing.