Investor Heterogeneity and Factor Pricing

Grace Xing Hu, Zhao Jin and Jianfeng Yu*

January 15, 2025

ABSTRACT

This paper shows that institutional ownership plays an important role in the pricing of a broad set of factors. Characteristic (Macro-related) factors earn a significantly smaller (larger) risk premium by 72 (54) basis points per month within stocks held by institutions compared to those held by retail investors. This empirical evidence suggest that the "flat" risk premium puzzle documented in the existing literature can be partly attributed to the tug-of-war between institutional and retail investors. Among institutional investors, investment companies and advisors, especially those excluding mutual funds and predominantly comprising hedge funds, contribute the most to these findings. Institutional investors, being relatively more rational, demand a premium for bearing systematic risks. Thus, the findings indicate that macro-related factor are proxy for fundamental risks, whereas characteristic factors are proxy for mispricings.

Key Words: Factor Pricing, Institutions, Investor Heterogeneity

JEL classification: G12, G20

*Hu (hux@pbcsf.tsinghua.edu.cn) and Yu (yujf@pbcsf.tsinghua.edu.cn) are from PBC School of Finance at Tsinghua University. Jin (zhaojin@cufe.edu.cn) is from School of Finance at Central University of Finance and Economics. We thank Jason Huafeng Chen, Darrell Duffie, Bing Han, Jun Pan, S. Ghon Rhee (Session Chair), Dacheng Xiu, and seminar participants at PBC School of Finance at Tsinghua University, School of Finance at Central University of Finance and Economics, School of Finance at Zhongnan University of Economics and Law, Lingnan College at Sun Yat-sen University, the 7th International Young Finance Scholars Conference, and the 2021 Korean Economic Review International Conference for their helpful comments. All errors are our own.

1 Introduction

Existing research find that beta-sorted portfolios for many "priced" factors discovered in the finance literature, including the most prominent market factor, exhibit negligible and often insignificant return spreads. This phenomenon, known in the literature as the "factor pricing failure", presents a significant challenge to the extensive body of research on factors and their pricing implications.[1] Several recent studies attempt to explain the puzzle in the time-series dimension, showing that the performance of factor models varies over time and is influenced by conditions such as money illusion (Cohen et al., 2005), macroeconomic announcements (Savor and Wilson, 2014), leverage aversion (Frazzini and Pedersen, 2014), aggregate disagreement (Hong and Sraer, 2016), investor sentiment (Shen et al., 2017), and day and night trading sessions (Hendershott et al., 2020).

In this paper, we examine the factor pricing failure from the cross-sectional perspective. Our findings reveals that institutional ownership plays an important role in the pricing of a broad set of factors. In particular, we observe a stark contrast in the impact of institutional ownership on two different categories of factors: (1) characteristic factors, which are portfolios sorted based on firm characteristics, such as the size and value factors documented in Fama and French (1993); and (2) macro-related factors, which are theoretically ground factors as in the consumption-based CAPM (Lucas, 1978; Breeden, 1979), CAPM (Sharpe, 1964), and production-based CAPM (Jermann, 1998; Belo, 2010), and innovations in pervasive macro-related state variables as in ICAPM (Merton et al., 1973), including variables like consumption growth, market returns, TFP growth, and inflation innovations.[2] This divergence underscores the complexity of factor pricing and helps shedding light on the distinct mechanisms that govern the pricing of characteristic and macro-related factors.

The key insights of this paper can be demonstrated through Figure 1. Stocks are double sorted into 5×5 groups independently based on their factor betas and levels of institutional ownership. Every data point within Figure 1 represents one of the 25 portfolios, with its

[1] Previous studies have documented and discussed about the phenomenon of insignificant pricing premiums, including Black (1972), Cohen et al. (2005), Frazzini and Pedersen (2014), Shen et al. (2017), Chen et al. (2020a), Daniel et al. (2020b), among many others.

[2] The factors examined in this paper are reported in Table 2 and further discussed in Appendix A.

average monthly excess returns on the y-axis (reported in percentages) and its respective factor betas on the x-axis. To allow for an equal-footing comparison across different types of factors, we compare the average excess returns against the composite betas – the average rankings of each portfolio's factor betas – calculated separately for the 12 characteristic factors in Panel A and for the 10 macro-related factors in Panel B. The levels of institutional ownership, denoted as RIO in the paper, are size-adjusted ratios of institutional holdings to mitigate the influence of firm size on ownership levels. We have also added the implied security market lines, in dash lines, to provide a clear visual comparison of how pricing patterns shift with respect to institutional ownership levels.

Consistent with previous research on the failure of factor pricing, Figure 1 shows a flat relation between stock returns and their factor betas. Pooling all 25 portfolios together, there is no evident link between the portfolios' excess returns and their composite betas, for both characteristic and macro-related betas. Yet, this dynamic shifts when focusing on portfolios with different levels of institutional ownership. For characteristic factors, there is a noticeable upward beta-return trend within the five portfolios having the lowest RIO, whereas this trend remains absent in other portfolios. For macro-related factors, an inverse pattern emerges: a significant positive correlation between excess returns and portfolio betas is observable in the five portfolios with the highest RIO, with no clear pattern in the rest.

Based on the 5×5 double-sorted portfolios, we further quantify the shift in the factor premiums – defined as the return differences between high- and low- composite beta stock portfolios – across groups with different RIO levels. For characteristics-based factors, we observe a consistent decline in factor premiums as institutional ownership increases. Specifically, the premium decreases from 0.32% per month for portfolios within the highest RIO quintile to -0.41% per month for those in the lowest quintile, leading to a significant monthly premium differential of -0.72%. Conversely, for macro-related factors, the factor premium demonstrates a positive correlation with institutional ownership levels, increasing from -0.12% per month in the lowest RIO quintile to 0.41% per month in the highest quintile, resulting in a positive monthly premium differential of 0.54%.

The opposite slopes of factor premiums with respect to institutional ownership – negative for characteristic factors and positive for macro-related factors – remain robust across a variety of different estimation methods. We observe similar results regardless of whether portfolio returns are adjusted using the CAPM or Fama-French 3-, 4-, or 5-factor models, whether factor premiums are calculated from equal- or value-weighted portfolios returns, or whether factor betas are estimated through an alternative two-factor model including both the market factor and the testing factor. Additionally, our analysis shows consistent results across individual factor betas: 10 out of 12 characteristic factors exhibit a significant negative factor premium differential across the RIO quintiles, with an average differential of -0.36% per month; meanwhile, 9 out of 10 macro-related factors exhibit significant positive premium differentials, averaging 0.27% per month. The strong impact of institutional ownership on factor premiums is further confirmed by the standard Fama-Macbeth regression analysis, which shows that the interaction term between RIO and factor betas is negative for characteristic factors and positive for macro-related factors.

These patterns in factor premiums with respect to institutional ownership suggest distinct pricing mechanism for characteristic and macro-related factors. Institutional investors, known for their more rational trading behavior, demand premiums for bearing systematic risks – a tendency that is less pronounced among retail investors.[3] Macro-related factors, which fetch the highest risk premiums in stocks primarily owned by institutional investors, likely indicate systematic risks. Conversely, characteristic factors, which command the highest risk premiums in stocks predominantly held by retail investors – with minimal institutional ownership – suggest that these factors are more likely associated with mispricing rather than fundamental risks. This is because factor betas could serve as proxies for mispricing, and high-beta firms tend to be undervalued (Stambaugh and Yuan, 2017; Daniel et al., 2020a). As retail investor participation increases, this mispricing effect becomes more pronounced,

[3]For example, Locke and Mann (2005) find that professional traders exhibit rational and disciplined trading behaviors, in strong contrasts to individual investors. Barber and Odean (2008) find that retail investors' buying decisions are influenced by attention, while institutions are not. Ben-Rephael et al. (2017) find that institutional attention responds more quickly to major news events, leads retail attention, and facilitates permanent price adjustment.

leading to a larger return spread among firms with low RIO.[4] While we acknowledge that empirically distinguishing between mispricing and risk is inherently challenging (Kozak et al., 2018), the contrasting patterns in factor premiums underscore the difficulty of reconciling both characteristics and macro-related factors within a purely rational framework.

A detailed examination of various categories of institutional investors provides additional empirical support for our hypothesis. Following the literature, we classify institutional investors into several groups: investment companies and advisors, banks, insurance companies, public and private pensions, university endowments, and others. Our analysis shows that stocks with high ownership by investment companies and advisors experience a -0.74% per month lower premium for characteristic factors and a 0.60% per month higher premium for macro-related factors compared to stocks with low ownership. Further analysis within investment companies and advisors reveals that components excluding mutual funds – primarily hedge funds – are the main drivers of these effects, showing the largest and most significant impact on factor premiums. In contrast, mutual funds exhibit weaker effects, producing a -0.36% per month lower premium for characteristic factors and a 0.32% per month higher premium for macro-related factors. These effects are marginally statistically significant and are largely driven by non-index mutual funds, as opposed to index mutual funds. For the rest of the institution types, similar but weaker patterns are observed for banks and public pensions, while no significant impact is found for insurance companies, private pensions, university endowments, or others. Overall, the findings suggest that actively managed institutional investors, rather than passive ones, are the primary drivers of the divergent patterns in premiums between characteristic and macro-related factors.

Next, we examine how institutional ownership influences factor pricing over time. First, after controlling for variables known to relate to factor premiums, such as market sentiment, inflation fluctuations, broker-dealer leverage, and aggregate disagreement, the effect of institutional ownership on factor premiums remains strong and consistent throughout most of our sample period. Second, there is large variations in the effect of institutional ownership

[4]Here, we use the term "mispricing" in a broad sense, referring to mispricing induced by sentiment or limits to arbitrages (Stambaugh and Yuan, 2017; Daniel et al., 2020a), as well as other frictions, such as noise trading (Huang et al., forthcoming), that are unrelated to fundamentals.

over time, particularly in connection with time-varying market sentiment. Specifically, for characteristic factors, risk premiums in low RIO stocks are larger (0.98% per month) following periods of high sentiment, which are characterized by excessive investor participation and mispricing. In contrast, for macro-related factors, risk premiums in high RIO stocks tend to be higher (0.89% per month) after periods of low sentiment, when the market atmosphere is more calm and rational.[5]

Finally, we address two potential alternative explanations: varying levels of short-sale constraints and noise trading risks, both of which co-vary with stocks' institutional ownership and could potentially affect factor pricing. Using tripe-sorted portfolios to control for these risks, our findings remain robust. Interestingly, the negative factor premium differential for characteristics factors is more pronounced in stocks with higher short-sale constraints, while the positive factor premium differential for macro-related factors is stronger in stocks with lower short-sale constraints. Both factor premium differentials remain significant across stocks with varying levels of noise trading risks, with slightly larger effects in those with low noise trading risks. These findings not only confirm the robustness of our findings but also further support the view that characteristics factors proxy for mispricing, while macro-related factors reflect systematic risks.

Our paper contributes to the literature in multiple aspects. A large number of recent studies investigate the failure of CAPM and other factor models, including money illusion (Cohen et al., 2005), FOMC announcements (Savor and Wilson, 2014), leverage aversion (Frazzini and Pedersen, 2014), disagreement (Hong and Sraer, 2016), sentiment (Shen et al., 2017; Chen et al., 2020b), and even day and night (Hendershott et al., 2020). All of these studies examine the failure of factor pricing models through the lens of time-series analysis. Our paper proposes a new explanation from the cross-sectional perspective, focusing on the opposite impact of institutional ownership on the pricing of characteristic and macro-related factors. Moreover, most studies only examine on the pricing of the market factor, with very

[5] Stambaugh et al. (2012) and Shen et al. (2017) find that fluctuations in investor sentiment lead to changes in asset valuations, with assets typically undervalued during low sentiment periods and overvalued during high sentiment periods. Short-sale constraints hinder arbitrage against overvaluation but do not affect the correction of undervaluation. As a result, the market is relatively calmer and more rational during low sentiment period, but exhibits greater mispricing during high sentiment period.

few exceptions such as Shen et al. (2017). Here, we study a broad set of characteristics-based and macro-related factors, simultaneously.

Our paper also complements the large body of research on institutional investors. Institutions are often assumed to be better informed and more rational than retail investors. Despite this, their performance and market impact remain subjects of ongoing debate. Previous studies, such as those by Nagel (2005), Boehmer and Kelley (2009), and Calluzzo et al. (2019), have suggested that institutions can enhance market efficiency. Conversely, research by Gompers and Metrick (2001), Basak and Pavlova (2013), and Lewellen (2011) provides limited support for the stock-picking abilities of institutions, showing that their portfolios often resemble market portfolios. McLean et al. (2020) find that trades by various types of institutional investors do not consistently predict returns. Recent work by Koijen and Yogo (2019) develops a heterogeneous demand-based model and analyzes the market impact of institutional demand from a neutral standpoint. We contribute to this strand of literature by showing that institutional ownership, particularly by investment companies and advisors, have direct market impact on a broad set of factors.

Lastly, this paper adds to the ongoing debate about whether characteristic factors act as proxies for risk or mispricing. On one side of the debate, studies such as Fama and French (1993), Sadka (2006), and Gonçalves et al. (2020) attempt to rationalize characteristic factor premiums as compensation for certain types of risks. On the other side, many research supports the mispricing hypothesis, including Stambaugh et al. (2012), McLean and Pontiff (2016), Stambaugh and Yuan (2017), and Engelberg et al. (2018). Recent research have also highlighted potential non-traditional fundamental risks embedded in characteristic factors, such as noise trading risks (Huang et al., forthcoming). Our research adds evidences that characteristic factors are more likely to capture mispricing (or non-fundamental risks) by examining the contrasting patterns of factor premiums with respect to institutional ownership.

2 Data Description

In this section, we outline the data sources and methodologies used to construct the institutional ownership measure, along with the set of 12 characteristic and 10 macro-related factors. We then present evidence of factor pricing failure, demonstrating that none of the factor premiums, derived from the excess returns of long and short beta-sorted portfolios, are statistically or economically significant.

2.1 Institutional Ownership

The quarterly institutional ownership data is obtained from the WRDS Thomson Reuters Institutional (13f) Holdings - S34 database. The sample period ranges from 1980Q1 to 2021Q4. Common stocks from NYSE/AMEX/NASDAQ in the CRSP dataset are included, and stocks reported in CRSP without records in 13F are excluded from the sample. The 13F database includes only the long positions of institutions, which can result in reported institutional ownership exceeding 100%. To address this, we winsorize the institutional ownership (IO) measure at the 0.01% and 99.99% levels.

Institutional investors are often guided by mandates and face a variety of constraints when making investment decisions (Cao et al., 2017). Many studies, including Ferreira and Matos (2008), have documented a strong positive correlation between institutional holdings and firm size. To account for size effect, we follow the methodology of Nagel (2005), regressing institutional ownership (IO) on firm size and its squared term each quarter. The resulting residuals, referred to as residual institutional ownership (RIO), are used as a proxy for the portion of institutional holdings not driven by the size effect.

Specifically, we first map IO, which ranges from 0 to 1, to a real line by using a Logit transformation:

$$\text{logit}(\text{IO}_{i,t}) = ln(\frac{\text{IO}_{i,t}}{1-\text{IO}_{i,t}}) \tag{1}$$

We then perform a cross-sectional quarterly regression of $\text{logit}(\text{IO}_{i,t})$ on the logarithm of firm size $ln(\text{ME}_{i,t})$ and its squared term $ln^2(\text{ME}_{i,t})$. The time-series averages of the estimated

coefficients are reported below, with Newey-West adjusted t-statistics in parentheses:

$$\text{logit}(\text{IO}_{i,t}) = \underset{(-21.89)}{-7.36} + \underset{(14.34)}{1.80} \times ln(\text{ME}_{i,t}) \underset{(-11.14)}{-0.09} \times ln^2(\text{ME}_{i,t}) + \text{RIO}_{i,t} \tag{2}$$

Table 1 presents the summary statistics for institutional ownership (IO) and residual institutional ownership (RIO). The average IO have shown significant growth, increasing from 0.20 in the 1980s to 0.62 during the 2020-2021 period. Alongside this upward trend in IO, there has been an improvement in data coverage, with the proportion of stocks having institutional ownership data increasing from 80% in the 1980s to 95% post-2020. By construction, RIO maintains a mean of zero, representing the residuals derived from cross-sectional regressions. Both IO and RIO demonstrate considerable variability, exhibiting full-sample standard deviations of 0.31 and 1.90, respectively. Figure 2 illustrates the time series of the cross-sectional averages, as well as the first and third quartiles for both IO and RIO.

2.2 Factors

2.2.1 Factor Description

We examine a comprehensive set of factors from the literature, categorizing them into two groups: characteristic factors and macro-related factors. Characteristic factors are long-short portfolios based on firm characteristics that have strong explanatory power of cross-sectional expected returns. We consider 12 widely recognized characteristic factors, including the five Fama-French factors including size (SMB), book-to-market ratio (HML), investment (CMA), and operating profitability (RMW), the momentum factor (MOM), the q^5 factors based on investment (R_{IA}), return on equity (R_{ROE}), and expected investment growth (R_{EG}), the gross profitability factor (PMU) based on firms' gross profits, two mispricing factors derived from six characteristics on firms' management ability (MGMT) and five characteristics related to firms' performance (PERF), and the AQR's quality factor (QMJ) based on 23 characteristics measuring firms' overall qualities.

Macro-related factors, on the other hand, are innovations of macro-related variables that

could affect future investment opportunities. According to ICAPM model (Merton et al., 1973), these variables show up in the stochastic discount factors (SDF) of investors who make investment decisions and maximize utilities inter-temporarily. Since macro-related factors capture fundamental risks, they are expected, in theory, to demand risk premiums. We consider 10 widely-studied macro-related factors in this paper, including consumption growth (CON), total factor productivity growth (TFP), industrial production growth (IPG), changes of default spread (DEF), changes of term spread (TERM), unexpected inflation (UI), changes of expected inflation (DEI), market excess return (MKT), changes of market volatility (VOL), and labor income growth (LAB).

Table 2 reports the descriptions and corresponding reference for the 12 characteristic factors and 10 macro-related factors we focus in this paper. A detailed description of these factors can be found in Appendix A.

2.2.2 Factor Beta Estimation

Estimating factor premiums requires forming portfolios with high and low exposures (betas) to the relevant factors. To form beta-sorted portfolios, we first estimate pre-formation betas (denoted as β^{pre}) for individual stocks listed on NYSE, AMEX, and NASDAQ exchanges. For each individual stock, pre-formation betas are estimated using data from the prior 5 years (with a minimum of 24 months), covering the period from January 1980 to December 2021. Specifically, the pre-formation beta of stock i on factor k at month t, denoted as $\beta^{pre}_{i,k,t}$, is estimated based on following equation:

$$r_{i,\tau} - r_{f,\tau} = \alpha_{i,k,t} + \beta^{pre}_{i,k,t} \times f_{k,\tau} + \epsilon_{i,k,\tau}. \tag{3}$$

Here, τ represents the months within the five-year window preceding time t, $r_{i,\tau}$ is the return of stock i at month τ, $r_{f,\tau}$ is the risk-free rate at month τ, and $f_{k,\tau}$ is the value of factor k at month τ.

For macro-related factors, pro-cyclical ones are expected to exhibit positive factor premiums. Default spreads (DEF) and volatility (VOL), being counter-cyclical factors, typically

rise during economic downturns, and thus their factor premiums are expected to be negative, as high exposures to these factors provide a hedge against market downturns. For the two inflation factors, unanticipated inflation (UI) and changes in expected inflation (DEI), the expected signs are less clear. Following Shen et al. (2017), we treat inflation as a pro-cyclical variable and hypothesize that UI and DEI will have positive factor premiums. To ensure all macro-related factors exhibit positive premiums, we flip the signs of the βs for DEF and VOL in the subsequent analysis. Characteristic factors are designed to have positive premiums by going long on stocks with higher expected returns and shorting those with lower expected returns.

To reduce noise in individual factor beta estimates and better capture aggregate factor exposures, we construct composite factor scores (denoted as β^{comp}) for each stock. By aggregating single-factor exposures within the same category, we form portfolios based on composite characteristic or macro-related factor exposures, allowing us to estimate average factor premiums for each category. Specifically, the composite beta score is calculated as the average of cross-sectional percentile rankings of single-factor betas. Each month t, the pre-formation betas $\beta^{pre}_{i,k,t}$ for individual factors k are ranked across all stocks, producing percentile rankings $\beta^{rank}_{i,k,t}$ (with descending rankings applied for the two counter-cyclical macro-related factors). The composite beta score $\beta^{comp}_{i,j,t}$ for stock i in month t is then the average of the ranking percentiles for all characteristic or macro-related factors:

$$\beta^{comp}_{i,j,t} = \frac{1}{N_j} \sum_{k=1}^{N_j} \beta^{rank}_{i,k,t}. \tag{4}$$

Here, $j = 1$ denotes the characteristic composite beta and $j = 2$ denotes the macro-related composite beta. N_j is the total number of characteristic (macro-related) factors, with $N_1 = 12$ and $N_2 = 10$.

2.2.3 Failure of Factor Pricing

In this subsection, we demonstrate the failure of factor pricing by showing a flat relationship between returns and betas for the beta-sorted portfolios. At the end of each month t, stocks

are sorted into deciles based on their pre-formation beta ($\beta^{pre}_{i,k,t-1}$) for each of the 22 factors. To ensure data availability, only information up to month $t-1$ is used for portfolio formation at time t. High-minus-low portfolios are then constructed by going long on the highest beta deciles and shorting the lowest beta deciles. Portfolio returns are equal-weighted.[6]

For each of the formed portfolios, we estimate the post-formation beta (denoted as $\beta^{post}_{p,k}$) as the realized factor exposures over the entire sample period. Unlike $\beta^{pre}_{i,k,t}$, which is estimated using the 5-year data prior to portfolio formation, $\beta^{post}_{p,k}$ captures the realized factor exposures based on the full-sample returns of the formed portfolios. Specifically, the realized factor exposure of portfolio p on factor k is calculated as follows:

$$r_{p,t} - r_{f,t} = \alpha_{p,k} + \beta^{post}_{p,k} \times f_{k,t} + \epsilon_{p,k,t}. \tag{5}$$

Here, $r_{p,t}$ is the return of portfolio p at month t, $r_{f,t}$ is the risk-free rate at month t and $f_{k,t}$ is the value of factor k at month t.

Table 3 presents the excess returns and post-formation betas for the highest beta, lowest beta, and the high-minus-low long-short portfolios constructed based on factor loadings. The results clearly illustrate the factor pricing failure. Across all 22 factors, none of the high-minus-low portfolios exhibit significantly positive excess returns, despite having large and statistically significant post-formation betas.

For characteristic factors, the highest factor premium, associated with PMU, is just 0.41% per month, with a marginal t-statistic of 1.61. In fact, five of the long-short portfolios even exhibit negative returns, though these are not statistically significant. In contrasts to these insignificant factor premiums, all characteristics-based high-minus-low beta portfolios display large, positive, and significant post-formation betas. These contrasting observations cast doubt on the validity of firm characteristic as genuine risk factors because they fail to generate significant return spreads across firms with varying exposures to these factors,

[6]Following the recommendation of Nagel (2005), we report the equal-weighted returns throughout most of our analysis to minimize size-related effects. For consistency, we also report the results based on equal-weighted returns in this section. The failure of factor pricing remains robust when value-weighted returns are employed.

despite the fact that tradable factors sorted directly by characteristics have shown significant expected returns.

Similarly, for macro-related factors, none of the high-minus-low portfolios exhibit significant excess returns. For example, the market factor (MKT), widely used in models like CAPM and Fama-French, yields a modest premium of only 0.27% (t-stat = 0.80) for high-minus-low portfolios, despite having significant factor loadings. Consumption growth (CON), a factor with strong theoretical grounding in the CCAPM model, even shows a negative premium of -0.04% (t-stat = -0.24). We acknowledge that noise in beta estimation, particularly for macro-related factors like consumption growth, makes it challenging to estimate factor premiums from beta-sorted portfolios. Nevertheless, even when portfolios are formed using composite betas, factor premiums remain insignificant (-0.05% and 0.16%) for both characteristic and macro-related factors.

3 Baseline Empirical Results

In this section, we present our main findings on the impact of institutional holdings on factor premiums. Using both portfolio sorting and the Fama-MacBeth regression methods, we demonstrate that stocks with high institutional ownership exhibit lower premiums for characteristic factors but higher premiums for macro-related factors. Further analysis of institutional investor types reveals that active institutions, such as investment companies and hedge funds, are the primary drivers of these patterns. These findings underscore the important role of investor heterogeneity on factor premium dynamics.

3.1 Portfolio Analysis

To quantify the impact of institutional ownership on factor pricing dynamics, we construct double-sorted portfolios and estimate factor premiums for stocks with different levels of residual institutional ownership (RIO). At the end of every month t, stocks are independently double-sorted into 5×5 groups based on their residual institutional ownership levels

$\text{RIO}_{i,t-1}$ estimated from regression (2) and composite factor exposures $\beta^{comp}_{i,j,t-1}$ estimated from regression (4). For both sorting variables, we only use values up to $t-1$ to ensure data availability at the portfolio formation month t. Following Nagel (2005), our main analysis rely on equal-weighted portfolio returns to mitigate the size effect.[7]

Table 4 reports the excess returns and CAPM αs for double-sorted portfolios with the highest and lowest β^{comp} for each of the five RIO quintitles. High-minus-low portfolios are constructed by taking a long position in the highest β^{comp} portfolios and a short position in the lowest β^{comp} portfolios within each RIO quintitles. The excess returns of these high-minus-low long-short portfolios reflect factor premiums estimated from stocks with different RIO levels. The CAPM αs measure factor premiums after adjusting for market exposures.[8] Lastly, we also report the the factor premium differences between the highest and lowest RIO quintitles, which capture the impact of institutional ownership on factor pricing.

The results in Table 4 reveal significant differences in factor premiums between high and low RIO stocks, with contrasting patterns for characteristic factors and macro-related factors. For characteristic factors, the factor premiums are substantially lower in high RIO stocks compared to low RIO stocks. The excess returns of high-minus-low portfolios decline monotonically from 0.32%, 0.13%, 0.20%, -0.16% to -0.41% per month as we move from the lowest to the highest RIO groups. This downward trend remains robust after adjusting for market exposures. The high-minus-low CAPM α is 0.86% per month in the lowest RIO stocks, compared to -0.04% in the highest RIO stocks, resulting in a substantial gap of -0.90% per month (t-stat=-4.26). Conversely, for macro-related factors, the results display an opposite trend. Substantial differences in factor premiums between high and low RIO stocks are observed, but with increasing premiums for higher RIO levels. The excess returns of long-short portfolios increase monotonically from -0.12% to 0.41% per month as RIO levels

[7]The results are robust when value-weighted portfolios are used. The main results based on value-weighted and log value-weighted portfolios are reported in Table C1 in Appendix C.1.

[8]Selecting the appropriate factor model for adjusting risk exposures is challenging due to the absence of a universally accepted fundamental factor model. We report CAPM αs for two main reasons: first, the CAPM is the most theoretically grounded and widely used factor model; second, several characteristics analyzed, such as size and book-to-market ratio, are already incorporated in models such as the Fama-French factor models. Importantly, our results remain robust when controlling for exposures to other factor models, as detailed in Table C2 in Appendix C.2.

increase, leading to a significant factor premium gap of 0.54% (t-stat=2.63).[9] The high-minus-low CAPM α differences remain significant at 0.63% (t-stat=2.96), indicating a robust impact of institutional holdings on factor premiums after adjusting for market exposures.[10]

The contrasting effects of institutional ownership on characteristic and macro-related factors suggest distinct pricing mechanisms for these two factor types. Institutional investors are generally more rational, seeking compensation for systemic risks while exploiting mispricings through arbitrage. Retail investors, by contrast, tend to exhibit behavioral biases and are more vulnerable to mispricing.[11] The decline in characteristic factor premiums as institutional ownership increases indicates that these factors are likely driven by mispricing rather than fundamental risks. In high RIO stocks, institutional arbitrage reduces mispricings, resulting in negligible characteristic factor premiums: high-minus-low portfolios show an average excess return of -0.41% and a CAPM alpha of -0.04%, both insignificant. Conversely, in low RIO stocks, where mispricing is more pronounced, high-minus-low portfolios yield a positive average excess return of 0.32% and a significant CAPM alpha of 0.86%. On the other hand, macro-related factor premiums increase with institutional ownership, reflecting stronger connections to systematic risks. For high RIO stocks, beta-sorted high-minus-low portfolios deliver the highest average excess return, at 0.41% per month. In low RIO stocks, however, short-sale constraints exacerbate overpricing in high-beta stocks (Stambaugh et al., 2015), leading to lower future returns: high-minus-low portfolios in this group show a negative average excess return of -0.12% and a significant CAPM alpha of -0.61%.

In addition to portfolios sorted by composite betas, we analyze portfolios sorted by single-factor betas. Specifically, we double-sort stocks into 5×5 groups based on residual insti-

[9]Empirically, there is also substantial noise in the beta estimation for macro-related factors, such as consumption growth, making it challenging to observe significant factor premiums and isolate the impacts of RIO. Additionally, the marginally significant factor premiums can partly be attributed to the high variance over time, as discussed in the Section 4.

[10]The last column of return differences between high and low RIO stocks indicates the outperformance of high RIO stocks compared to low RIO stocks. These findings align with Gompers and Metrick (2001), who show the predictive power of institutional holdings on stock returns, and attribute it to demand shocks driven by institutional investors. Rather than focusing on the link between institutional holdings and future returns, this paper emphasizes the relationship between institutional ownership and factor premiums

[11]Extensive literature highlights the sophistication of institutional trading behavior compared to retail investors, who are more prone to behavioral biases and often regarded as noise traders; see Locke and Mann (2005); Barber and Odean (2008); Kumar (2009); Ben-Rephael et al. (2017).

tutional ownership (RIO) and single-factor betas for each factor. The portfolio formation follows the same method as for composite-beta portfolios, except we use individual factor betas for sorting at month t. For counter-cyclical macro-related factors (DEF and VOL), we reverse the signs of the betas during sorting. To mitigate measurement errors associated with single-factor betas, we also construct averaged portfolios for multiple factors within the same factor categories.[12]

The results remain robust when analyzing single-factor premiums. Table 5 reports the factor premiums for individual factors, measured as the excess returns of high-minus-low β portfolios within both high and low RIO groups, along with their differences. For characteristic factors, the premium differences are predominantly negative, supporting the idea that these factors have greater explanatory power in stocks with low institutional ownership. In the averaged portfolios, factor premiums in the low RIO group are consistently higher than in the high RIO group, resulting in significant differences of -0.36% (t-stat = -3.06), -0.43% (t-stat = -3.11), and -0.59% (t-stat = -3.60) per month. Among portfolios based on 12 single factors, 10 show larger premiums in low RIO stocks, with differences ranging from -0.22% for the Fama-French investment factor (CMA) to -0.83% for the quality factor (QMJ). The size factor (SMB) is an exception, showing a 0.40% higher premium in high RIO stocks. This suggests that SMB, one of the most widely used characteristic factors, may also capture systematic risks rather than solely reflecting mispricing, given that smaller firms tend to be riskier and more sensitive to macroeconomic conditions (Perez-Quiros and Timmermann, 2000). Consequently, the size factor aligns more closely with macro-related factors than with characteristic ones.

In contrast, macro-related factors exhibit significantly larger premiums in stocks with high institutional ownership. For the averaged portfolios, the excess returns of high-minus-low beta portfolios, which proxy for average factor premiums, are significantly higher in high

[12] Average portfolios are constructed as equal-weighted averages of portfolios based on single-factor betas. For characteristic factors, AVG 1-12 represents the average of portfolios based on all 12 characteristic factors, while AVG 2-12 excludes portfolios based on size. AVG 10-12 includes only portfolios based on composite characteristics such as MGMT, PERF, and QMJ. For macro-related factors, AVG 1-10 represents the average of portfolios based on all 10 macro-related factors. AVG 1-6 includes portfolios based on six factors with strong economic foundations, specifically CON, TFP, IPG, LAB, MKT, and VOL, while AVG 1-8 adds DEF and TERM to these six factors.

RIO stocks. The average factor premiums are 0.22%, 0.26%, and 0.24% in high RIO stocks, compared to -0.05%, -0.06%, and -0.06% in low RIO stocks, resulting in significant differences of 0.27% (t-stat = 3.45), 0.32% (t-stat = 3.18), and 0.30% (t-stat = 3.47) per month. Among the 10 single factors, 9 exhibit positive differences in factor premiums. Notably, the market excess returns (MKT) in the CAPM model show a difference of 0.49% per month (t-stat = 2.47), with higher premiums in high RIO stocks. These findings indicate that macro-related factors have greater explanatory power and larger premiums in stocks with high institutional ownership, suggesting that investor heterogeneity may help resolve puzzles surrounding factor pricing failures, particularly in the CAPM framework.

In summary, results from composite, averaged, and single-factor sorted portfolios consistently demonstrate that institutional ownership strongly influences factor pricing. Characteristic factors exhibit lower premiums in stocks with high institutional ownership, while macro-related factors show higher premiums in the same context.

3.2 Fama-MacBeth Regression

In this subsection, we use Fama-MacBeth regressions to analyze the impact of institutional ownership on factor pricing. Standard Fama-Macbeth regressions capture factor pricing by examining the cross-sectional beta-return dynamics, represented by the expected excess returns linked to factor exposures. To investigate the impact of institutional ownership, we extend the standard Fama-Macbeth regressions with an interaction term of residual institutional ownership (RIO) and factor exposures (β). The Fama-MacBeth regressions proceed in two steps. First, we run monthly cross-sectional regressions of excess stock returns on post-formation betas, RIO, and their interaction term:

$$r_{i,t+1} - r_{f,t+1} \;=\; a_{k,t} + b_{1,k,t}\ \beta^{post}_{i,k,t} \times \mathrm{RIO}_{i,t} + b_{2,k,t}\ \beta^{post}_{i,k,t} + b_{3,k,t}\ \mathrm{RIO}_{i,t} + \epsilon_{i,k,t+1} \tag{6}$$

Here, $\mathrm{RIO}_{i,t}$ is the residual institutional ownership level of stock i in month t, as defined in equation (2). For each individual factor k, we assign the post-formation betas of beta-

sorted portfolios to the constituent stocks within those portfolios ($\beta_{i,k,t}^{post}$) to minimize measurement errors. For counter-cyclical factors (DEF and VOL), we reverse the signs of the post-formation betas. The composite post-formation betas are calculated as the average percentile rankings of single-factor post-formation betas, with one set corresponding to characteristic factors and the other to macro-related factors. To ensure a fair comparison across different factors, both RIO and post-formation betas are standardized to have a cross-sectional mean of 0 and a standard deviation of 1 prior to forming the interaction terms. These interaction terms are subsequently calculated as the product of the standardized RIO and standardized betas.[13]

Table 6 presents the effects of RIO on factor beta-return dynamics by reporting the time-series averages of the coefficients in regression (6). The coefficient b_1 for the interaction term $\beta \times$ RIO measures the incremental impact of RIO on the relationship between beta and returns. The results show that the average b_1 is significantly negative for most characteristic factors and positive for most macro-related factors. Specifically, a one standard deviation increase in RIO weakens the beta-return relationship for 11 out of 12 characteristic factors, reducing monthly excess returns by an average of 0.11% for each standard deviation increase in beta. Conversely, for 9 out of 11 macro-related factors, higher RIO strengthens the beta-return relationship, leading to an average increase of 0.11% in monthly excess returns per standard deviation increase in beta. These effects are substantial compared to the unconditional beta-return relationships (measured by b_2), which are much smaller and statistically insignificant, averaging just 0.01% for characteristic factors and 0.06% for macro-related factors.

Overall, these findings confirm that RIO significantly influences factor pricing. Characteristic factors exhibit weaker beta-return relationships in stocks with higher institutional holdings, while macro-related factors show stronger relationships. This underscores the crucial role of investor heterogeneity in shaping factor pricing mechanisms and highlights the

[13]The results remain robust even after controlling for other firm characteristics. In Appendix C.4, we directly examine the impacts of institutional ownership (IO) while controlling for factors such as firm size, book-to-market ratio, turnover, volatility, and past returns. The findings, presented in Table C4, indicate a consistent negative (positive) impact of institutional ownership on characteristic (macro-related) factor beta-return relationships.

distinct pricing dynamics of characteristic and macro-related factors.

3.3 Different Types of Institutions

To further investigate how investor heterogeneity influences factor premiums, we examine the effects of holdings by different types of institutions. Prior studies suggest that active institutions, such as investment firms and hedge funds, are more adept at exploiting market mispricings, whereas passive institutions are less likely to engage in such strategies.[14]

For institutional investor types, we rely on the Institutional Investor Classification Data provided by Brian Bushee, which adjusts for errors in institutional type codes provided by Thomson Reuters.[15] We classify institutions into seven types: investment companies and advisors, bank trusts, insurance companies, public pension funds, private (coporate) pension funds, university and foundation endowments, and others.[16] Investment companies and advisors are further divided into three subtypes: index mutual funds, non-index mutual funds, and other investment companies (e.g., hedge funds), using data from the Thomson Reuters Mutual Fund Holdings (S12) Database. Mutual funds without investment objective codes (IOCs) are classified as index mutual funds. Following Koijen and Yogo (2019), the remaining funds, after excluding mutual funds, fall into the subtype of other investment companies, which are presumably dominated by hedge funds.[17]

[14] Ke and Ramalingegowda (2005) highlight the role of transient institutions in exploiting post-earnings announcement drift (PEAD), in contrast to relatively passive, long-term institutions. Calluzzo et al. (2019) show that active institutions, such as investment firms and advisors, increase anomaly positions following academic publications. Within investment companies, Akbas et al. (2015) find that mutual fund flows tend to exacerbate market mispricing, whereas hedge fund flows appear to correct anomalies and reduce aggregate mispricing.

[15] Data are downloaded from Brian Bushee's website. In Brian Bushee's classification, reliable Thomson Reuters type codes for institutions that existed before 1998 are carried forward in time. For new institutions, the type codes are assigned based on searches for information about the fund manager.

[16] The data originally classifies institutions into eight types. We aggregate investment companies and independent investment advisers into one group, termed "investment company and advisor", as suggested by Bushee.

[17] The S12 database assigns investment objective codes (IOC) to each mutual fund. The total assets of funds with missing IOCs surged from 0 in 1997 to more than half of the total mutual fund assets by 2021, mirroring the dramatic growth pattern of index funds. Additionally, the largest fund missing an IOC is the widely known SPY ETF, followed by various index funds offered by major financial institutions such as Vanguard and BlackRock. Therefore, we classify mutual funds without IOCs as index mutual funds and those with IOCs as non-index funds. Appendix B provides detailed information on mutual fund IOC categories and a list of the top 10 largest funds missing IOCs.

Figure 3 illustrates the annual market capitalization shares of institutional types from 1981 onward.[18] Investment companies and advisors grew significantly, increasing from 11.4% in 1981 to 49.2% in 2021, which is largely driven by the rapid rise of index mutual funds since the 2000s. By contrast, bank holdings decrease from 15.1% in 1981 to 11.1% in 20201. Insurance companies, private and public pension funds, and university endowments account for small and relatively unchanged shares.

We calculate residual institutional ownership (RIO) for each investor type using equation (2), replacing the IO variable with holdings specific to each institution type. Stocks are double-sorted into 5×5 groups based on their RIO and composite factor exposures, as described in section 3.1. Table 7 presents factor premiums for high- and low-RIO stocks across institutional types.

The results indicate that investment companies and advisors play a dominant role in driving the observed factor premium differences. For stocks with higher holdings by these institutions, characteristic factor premiums are significantly lower (-0.74%, t-stat = -3.01), while macro-related factor premiums are significantly higher (0.60%, t-stat = 2.84). Further analysis shows that other investment companies (excluding mutual funds and primarily hedge funds) are the primary contributors, with the largest observed differences: characteristic factor premiums decrease by -0.77% (t-stat = -3.26), and macro-related factor premiums increase by 0.66% (t-stat = 3.20). In contrast, mutual funds exhibit weaker effects, with characteristic factor premiums reduced by -0.36% (t-stat = -1.72) and macro-related factor premiums increased by 0.32% (t-stat = 1.65). Notably, the effects linked to mutual funds are primarily driven by non-index funds, while index mutual funds show no significant impact.

Among other institutional types, public pensions and banks exhibit similar effects on factor premiums. Higher holdings by public pension funds reduce characteristic factor premiums by 0.53% and increase macro-related factor premiums by 0.40%.[19] Similarly, higher

[18] Bank holdings in 1980 are significantly lower than those in subsequent years and are thus not considered reliable. Therefore, the analysis of institutional decompositions starts from 1981. The following sorting results remain robust when including the data from 1980.

[19] These findings align with previous research, which shows that public pension funds are more active in investment (Coggin et al., 1993; Artiga González et al., 2020) and corporate governance (Guercio and Hawkins, 1999) compared to private pensions, insurance companies, and banks.

holdings by banks lower characteristic factor premiums by 0.32% and increase macro-related factor premiums by 0.29%, with marginal significance. In contrast, no significant impact on factor premiums is observed for holdings by insurance companies, private pensions, university endowments, or other institutions.

Overall, the results confirm that investor heterogeneity is a key determinant of factor pricing. Active institutions, particularly hedge funds, show the strongest impact, while more passive institutions have a comparatively weaker impact.

4 Time Series Analysis

In this section, we examine how the effects of investor heterogeneity on factor premiums evolve over time. First, we demonstrate that the variations in factor premiums linked to institutional ownership remain robust throughout the entire sample period. Next, we compare the cross-sectional effects of institutional holdings on factor premiums with the time-series effects discussed in prior research. We find that our results remain robust even after accounting for these time-series effects. Furthermore, incorporating both cross-sectional and time-series effects reveals a stronger combined effects that can better explain variations in factor premiums, shedding new light on the puzzle of factor pricing failure.

4.1 Time-Series Variations in the Effects of Investor Heterogeneity

We start by examining the time-series variations of the effects of investor heterogeneity on factor premiums. In Figure 4, we present the time series of factor premiums estimated for stocks with high and low residual institutional ownership (RIO), along with their differences. The top panels illustrate the factor premiums for characteristic and macro-related factors in high (solid lines) and low (dashed lines) RIO stocks, while the bottom panels depict the differences between these groups. To minimize noise in short-term returns, we employ an exponential weighted moving average (EWMA) model with a decay parameter of 0.95.

The results in panels (a) and (b) reveal significant time-series variations in factor pre-

miums for both characteristic and macro-related factors across high and low RIO stocks. As discussed further in the next subsection, these fluctuations can be partially attributed to time-series variables influencing factor premiums, such as sentiment Shen et al. (2017) and inflation Cohen et al. (2005) documented in prior research. Given the paper's focus on examining the cross-sectional effects of institutional ownership, we concentrate on the differences between high and low RIO stocks, shown in Panels (c) and (d). The plots reveal stable and persistent differences in factor premiums between these groups. With few exceptions – brief periods around 2000, 2020, and 2021 – the high-minus-low factor premiums consistently remain below zero for characteristic factors and above zero for macro-related factors. These patterns underscore a robust impact of institutional ownership on factor premiums throughout the sample period.

4.2 Controlling for Time-Series Forces

Previous studies have shown that factor premiums are influenced by various time-series forces. For example, characteristic factors tend to exhibit higher premiums following periods of high sentiment (Chen et al., 2020b), whereas macro-related factors show lower premiums during these periods (Shen et al., 2017). Regarding the well-studied CAPM model, prior research suggests that the market factor is more significantly priced after periods of low inflation (Cohen et al., 2005), relaxed leverage constraints (Frazzini and Pedersen, 2014), or low disagreement among investors (Hong and Sraer, 2016). To investigate whether the cross-sectional effects of institutional ownership remain robust after controlling for these time-series variables, we directly compare their impacts using the following pooled regressions:

$$\begin{aligned} r_{i,t+1} - r_{f,t+1} \;=\; & a_j + b_{1,j}\ \beta_{i,j,t}^{comp_post} \times \mathrm{RIO}_{i,t} + b_{2,j}\ \beta_{i,j,t}^{comp_post} \times tsVar_t \\ & + c_{1,j}\ \beta_{i,j,t}^{comp_post} + c_{2,j}\ \mathrm{RIO}_{i,t} + c_{3,j}\ tsVar_t + \epsilon_{i,j,t+1} \end{aligned} \tag{7}$$

Here, $r_{i,t+1} - r_{f,t+1}$ represents the excess return of stock i in month $t+1$. $\beta_{i,j,t}^{comp_post}$ is the composite post-formation beta for stock i in month t for factor group j, where $j = 1$ denotes

the characteristic composite beta, and $j = 2$ denotes the macro-related composite beta. $\text{RIO}_{i,t}$ is the residual institutional ownership of stock i in month t. $tsVar_t$ represents the time-series variables, including investor sentiment, inflation changes, broker-dealer leverage, and investor disagreement.[20] $\beta_{i,j,t}^{comp_post}$ and $\text{RIO}_{i,t}$ are standardized to have a mean of 0 and a standard deviation of 1 across all stocks for each month, while $tsVar_t$ is standardized across time to have a mean of 0 and a standard deviation of 1.

Table 8 reports the regression results. For both characteristic and macro-related factors, the cross-sectional effects of RIO on beta-return relations remains robust after controlling for time-series variables. Specifically, an increase in RIO by one standard deviation weakens beta-return relations for characteristic factors, with monthly excess returns decreasing by 0.13% to 0.15% for a one standard deviation increase in beta. Conversely, for macro-related factors, higher RIO strengthens beta-return relations, with excess returns increasing by 0.11% to 0.13%.

The time-series variables also exert significant effects on beta-return relations, consistent with prior research. For characteristic factors, beta-return relations are strengthened during periods of high sentiment (coef of $\beta \times \text{Sent} = 0.40$), high inflation (coef of $\beta \times \Delta\text{Inf} = 0.16$), high leverage (coef of $\beta \times \text{Leverage} = 0.35$), or high disagreement (coef of $\beta \times \text{Disagree} = 0.08$). Conversely, macro-related factors exhibit weaker beta-return relations during these same periods, with interaction term coefficients dropping to -0.32, -0.16, -0.27, and -0.08, respectively.

In summary, these results not only validate prior research on the effects of time-series variables on factor pricing but also underscore the distinct and robust role of institutional ownership as a key determinant of factor premiums in the cross-sectional dimension.

[20]Monthly sentiment data is sourced from Jeffrey Wurgler's website. Changes in monthly inflation are calculated using Consumer Price Index (CPI) data from the Federal Reserve Bank of St. Louis FRED database. The broker-dealer leverage measure is defined as per Adrian et al. (2014) and obtained from Tyler Muir's website, with data available up to the third quarter of 2017. Investor disagreement, as defined by Yu (2011), is estimated using the aggregate standard deviations of long-term earnings-per-share forecasts from the IBES database, with data starting from December 1981.

4.3 Combined Effects of Cross-sectional and Time-series Forces

Next, we investigate the combined effects of cross-sectional and time-series influences on factor premiums. Considering the significant role of sentiment (Baker and Wurgler, 2006), our analysis focuses on how cross-sectional institutional ownership and time-series investor sentiment jointly impact factor premiums. In Appendix C.5, we provide additional discussions on the interplay between institutional ownership and other time-series variables, such as inflation, leverage, and disagreement.

Investor sentiment reflects (over)participation in the market and captures fluctuations in market rationality over time. Previous studies have established the role of sentiment in shaping factor premiums. For example, Antoniou et al. (2016) show that the security market line is upward-sloping after periods of low sentiment but downward-sloping after periods of high sentiment. Building on this, Shen et al. (2017) and Chen et al. (2020b) document opposing time-series patterns for characteristic and macro-related factors: characteristic factors are negatively priced following low sentiment periods but positively priced after high sentiment periods, whereas macro-related factors display the opposite trend.

Our objective is to examine the joint impact of institutional holdings, representing the cross-sectional dimension, and investor sentiment, representing the time-series dimension, on factor pricing. Since variations in rationality across stocks (cross-sectional) and over time (time-series) can both influence the strength of factor pricing, we hypothesize that when rationality is at its lowest (or highest) in both dimensions, the pricing power of characteristic (or macro-related) factors will be maximized, yielding the most prominent factor premiums.

To test this hypothesis, we analyze factor premiums within stocks grouped by RIO levels and across different sentiment periods. At the end of each month t, stocks are double-sorted into 5×5 groups based on their residual institutional ownership (RIO) and composite factor exposures. Factor premiums are calculated as the excess returns of high-minus-low beta-sorted portfolios within each RIO tier. The sample is split into high and low sentiment periods, determined by the median level of investor sentiment over the full sample. We then report average high-minus-low portfolio returns to examine the combined effects of RIO and

sentiment on factor premiums.

Table 9 summarizes the factor premiums across RIO levels and sentiment periods, along with their differences. The results show that characteristic factor premiums are highest (0.98% per month, t-stat = 2.00) when both cross-sectional and time-series rationality are low – that is, during high sentiment periods within the lowest RIO stocks. Conversely, macro-related factor premiums peak at 0.89% per month (t-stat = 2.33) when rationality is high in both dimensions – during low sentiment periods within the highest RIO stocks. A trading strategy that considers variations in both dimensions yields an average excess return of 2.01% (0.98% - (-1.03%)) per month for characteristic factors and 1.61% (0.89% - (-0.72%)) per month for macro-related factors – nearly three times the magnitude of the factor premium differences attributed solely to RIO differences in prior analyses.

Additionally, the factor premium differences attributable to either RIO or sentiment remain robust and significant when controlling for the other dimension. Across both high and low sentiment periods, the differences in factor premiums between high and low RIO stocks are consistently negative for characteristic factors (-0.77% and -0.68%) and positive for macro-related factors (0.66% and 0.41%). Similarly, within each RIO group, the factor premium differences between high and low sentiment periods are consistently positive (1.23% to 1.52%) for characteristic factors and negative (-0.60% to -1.20%) for macro-related factors.

In conclusion, our findings demonstrate that both cross-sectional (RIO) and time-series (sentiment) forces independently and jointly contribute to factor pricing. By combining these dimensions, we can significantly improve the explanatory power of factor pricing.

5 Alternative Interpretations

In this section, we consider two alternative explanations for our findings: short-sale constraints and noise trading risks. Stocks facing higher short-sale constraints or greater exposure to noise trading risks are more likely to encounter significant limits to arbitrage, which increases the potential for mispricing. Consequently, factors driven by mispricing may exhibit higher premiums, while the pricing of fundamental factors could be distorted. Given

that stocks with high institutional ownership are typically less affected by short-sale constraints and noise trading effects, one might argue that the observed factor pricing patterns, especially for characteristic factors, are attributable to these constraints and risks rather than institutional ownership. To address these alternative explanations, we control for short-sale constraints and noise trading risks using direct measures constructed from stock lending market data and mutual fund holdings. Our results show that the impact of institutional ownership on factor premiums remains robust even after controlling for short-sale constraints and noise trading risks, indicating that its influence extends beyond these alternative explanations.

5.1 Short Sale Constraints

In this section, we aim to examine whether short-sale constraints serve as the primary mechanism through which institutional ownership affects factor premiums by directly controlling for these constraints. If institutional ownership influences factor premiums solely through the channel of short-sale constraints, its effect should disappear once we account for direct measures of short-sale constraints. Conversely, if the observed impact of institutional ownership remains robust even after controlling for these constraints, it would suggest that its influence is driven by the more rational trading behaviors associated with institutional investors. Furthermore, since stocks with high short-sale constraints are more exposed to mispricing, we anticipate an interactive effect: the impact of RIO on characteristic factor premiums should be stronger for stocks with high short-sale constraints, while its effect on macro-related factor premiums should be more pronounced for stocks with low short-sale constraints.

To test this, we analyze three measures of short-sale constraints. The first measure, short interest, is the ratio of shares held short to total outstanding shares, as proposed by McLean et al. (2020), with data obtained from the Compustat Supplemental Short Interest File. The second measure, utilization, represents the proportion of a stock's value on loan relative to its total lendable value in the borrowing market. The third measure, indicative fee, reflects the borrowing cost of equities in the borrowing market. Data for utilization and indicative

fees are obtained from Markit Securities Finance Analytics, as detailed by Muravyev et al. (2022). The short interest data spans January 1980 to December 2021, while utilization and indicative fee data are available from May 2002 onward. In our analysis, higher short interest, lower utilization, and lower indicative fees indicate lower short-sale constraints.

Table 10 presents summary statistics for the three measures of short-sale constraints. On average, 2.80% of a stock's total outstanding shares are held in short positions. In the borrowing market, 15.48% of the equities provided by lenders are on loan, and the average borrowing cost is 2.50%. The relationships between short interest and indicative fee with institutional ownership (IO) and residual institutional ownership (RIO) are consistent with prior studies: higher IO or RIO is generally associated with higher short interest and lower indicative fees, indicating fewer short-sale constraints. In contrast, utilization exhibits no consistent correlation with institutional ownership, showing small magnitudes and opposite signs in its correlations with IO and RIO.

To address potential effects of short-sale constraints, we construct triple-sorted portfolios based on composite beta, residual institutional ownership (RIO), and one of the short-sale constraint (SSC) measures. This approach enables us to evaluate the impact of institutional ownership on factor premiums while controlling for short-sale constraints. To mitigate the issue of empty groups in triple sorting, we employ dependent sorting. Specifically, at the end of each month t, stocks are sequentially sorted into $2 \times 2 \times 5$ groups based on one of the SSC measures, RIO level, and composite factor exposures. To ensure data availability at the time of portfolio formation, only data up to month $t-1$ is used (one-month lag). If the effects of institutional holdings are entirely driven by short-sale constraints, no significant differences in factor premiums should remain across RIO groups after controlling for short-sale constraint levels.

Table 11 reports the factor premiums and their differences across high and low RIO groups after controlling for short-sale constraint levels, proxied by short interest, utilization, or indicative fee. For stocks within the same short-sale constraint level, we calculate the excess returns of high-minus-low composite beta portfolios within each RIO group as the

measure of factor premiums. Additionally, the differences in factor premiums and their CAPM alphas between high and low RIO groups are presented, highlighting the effects of institutional holdings after accounting for short-sale constraints.

The results clearly show that institutional holdings significantly influence factor premiums, even after controlling for short-sale constraints. For characteristic factors, factor premiums are consistently lower in high RIO stocks compared to low RIO stocks, especially in stocks with high short-sale constraints where mispricing is more difficult to arbitrage away. In such stocks, the differences in factor premiums between high and low RIO groups are -0.41%, -0.44%, and -0.71%, with CAPM alphas reaching -0.53%, -0.55%, and -0.87%, all statistically significant. For stocks with low short-sale constraints, while the results are less pronounced due to these stocks being less prone to mispricing, the factor premium differences remain largely negative, ranging from 0.01% to -0.50%, with CAPM alphas between -0.12% and -0.65%.

For macro-related factors, the factor premiums between high and low RIO stocks remain positive, particularly in stocks with low short-sale constraints, where fundamental risks predominantly drive expected returns. CAPM alphas are 0.18%, 0.29%, and 0.28% higher in high RIO groups for stocks with high short-sale constraints and 0.33%, 0.39%, and 0.07% higher for stocks with low short-sale constraints. These findings suggest that short-sale constraints, at least as measured in this study, cannot fully explain the observed effects of institutional ownership on factor premiums.

Furthermore, short-sale constraints appear to amplify the effects of RIO on factor premiums, with the direction of the interaction dependent on the type of factor. The impact of RIO on characteristic factors is more pronounced in stocks with high short-sale constraints, whereas its impact on macro-related factors is more pronounced in stocks with low short-sale constraints. This distinction further underscores the differing nature of the pricing mechanisms for characteristic and macro-related factors.

5.2 Noise Trading Risk

Previous literature demonstrates that non-fundamental demand shocks from noise traders can significantly affect asset prices and risk-return relationships (Long et al., 1990; Lou, 2012). Specifically, Huang et al. (forthcoming) suggests that noise trading risks are a key driver of characteristic factor premiums. Retail investors, compared to institutions, are more constrained and frequently act as noise traders (Grinblatt and Keloharju, 2009; Kumar, 2009). Consequently, stocks predominantly held by retail investors are more exposed to noise trading risks compared to those held by institutions, which may help explain the observed impact of institutional holdings on factor premiums. In this section, we directly control for noise trading risk exposures and shocks faced by each stock to address this alternative explanation.

Noise trading shocks are often measured using mutual fund flow-induced trading (FIT). We construct the FIT measure following Lou (2012). Specifically, we first calculate the flow of each mutual fund k in every quarter t as the percentage change in total net assets (TNA) after accounting for fund returns, as shown in Equation (8):

$$\text{flow}_{k,t} = \frac{\text{TNA}_{k,t} - \text{TNA}_{k,t-1} \times (1 + \text{RET}_{k,t}) - \text{MGN}_{k,t}}{\text{TNA}_{k,t-1}} \tag{8}$$

where $\text{MGN}_{k,t}$ represents the increase in TNA due to fund mergers in quarter t.

The FIT measure for each stock i is calculated as the last quarter's fund holding weighted average of fund flows, adjusted by the partial scaling factor (PSF), which quantifies the proportion of holdings transferred in response to the previous quarter's inflows or outflows:

$$\text{FIT}_{i,t} = \frac{\sum_k \text{Shares}_{k,i,t-1} \times \text{flow}_{k,t} \times \text{PSF}_{k,t}}{\sum_k \text{Shares}_{k,i,t-1}} \tag{9}$$

where $\text{Shares}_{k,i,t-1}$ represents the split-adjusted shares of stock i held by mutual fund k in quarter $t-1$. Following Lou (2012) and Huang et al. (forthcoming), we set $\text{PSF}_{k,t}$ to 0.970 for outflows and 0.858 for inflows.

Our data is obtained from the CRSP Mutual Fund database and the Thomson Reuters S12 mutual fund holdings database. We use the MFLinks from WRDS to merge mutual

funds across these databases. Equity funds are selected by requiring the fund objective code to begin with "ED" and by ensuring that common stock investments constitute between 75% and 105% of the portfolio. We exclude fund-quarter observations with quarter-end TNA below $1 million. The final sample includes 6,986 U.S. domestic equity mutual funds and 17,836 common stocks, resulting in 34,834,716 fund-stock-quarter observations spanning 1980Q1 to 2021Q4.

We measure stock-level noise trading risk exposure using the beta of an individual stock's FIT relative to the aggregate market FIT, where the market FIT is calculated as the value-weighted FIT of all individual stocks in each quarter. Specifically, we estimate the noise trading risk exposure, $\beta_{i,t}^{\text{FIT}}$, for stock i in quarter t using 8-year rolling window quarterly regressions, requiring a minimum of 20 quarters of data. The estimation is based on the following equation:

$$\text{FIT}_{i,\tau} = a_{i,t} + \beta_{i,t}^{\text{FIT}} \times \text{FIT}_{mkt,\tau} + \epsilon_{i,\tau} \tag{10}$$

where τ denotes the quarters within the eight-year window preceding time t. A higher $\beta_{i,t}^{\text{FIT}}$ indicates stronger co-movement with aggregate non-fundamental demand shocks, indicating greater exposure to noise trading risks.

Table 12 reports the factor premiums and their differences across high and low RIO groups after controlling for noise trading risk exposures (Panel A) or noise trading shocks (Panel B). At the end of each quarter t, stocks are triple-sorted into $2 \times 2 \times 5$ groups sequentially based on FIT beta or FIT level, residual institutional ownership (RIO) level, and composite factor exposures. For each group of stocks with the same levels of FIT and RIO, we report the factor premiums as the excess return differences between the highest and lowest composite beta stocks. Additionally, we present the differences in factor premiums and their CAPM alphas between high and low RIO groups, highlighting the effects of institutional holdings after accounting for noise trading risks or shocks.

The impact of institutional ownership on factor pricing remains robust and statistically significant, after accounting for noise trading risk exposures and shocks. When controlling

for noise trading risk exposures, the differences in characteristic factor premiums between high and low RIO groups remain negative, with risk-adjusted CAPM alphas of -0.88% (t-stat= -1.66) for the high FIT beta group and -1.30% (t-stat= -2.75) for the low FIT beta group. Macro-related factor premiums, on the other hand, show positive differences in risk-adjusted CAPM alphas, with 0.90% (t-stat= 1.77) for the high FIT beta group and 1.28% (t-stat=2.20) for the low FIT beta group. Similarly, after accounting for noise trading shocks, the differences in characteristic (macro-related) factor premiums between high and low RIO groups remain significantly negative (positive). These findings indicate that the observed effects of institutional ownership cannot be attributed to variations in noise trading exposures or shocks, underscoring the critical role of institutional ownership in shaping factor premiums.

6 Conclusion

This paper documents the significant impact of institutional ownership on the pricing of a wide range of factors. Using individual stocks as testing assets, we provide evidence that the flat risk premium puzzle documented in the literature can be partially attributed to the competing behaviors of institutional and retail investors. Specifically, we reveal a contrasting pattern in factor pricing across stocks: characteristic factors exhibit lower premiums in stocks with high residual institutional ownership, while macro-related factors show higher premiums. Furthermore, by categorizing institutions into different types, we identify investment companies and advisors – particularly those dominated by hedge funds after excluding mutual funds – are the primary contributors of these results. Our findings suggest that macro-related factors likely capture fundamental economic risks, considering institutional investors generally behave more rationally than retail investors and demand a premium for bearing systematic risks. By comparison, characteristic factors appear to be more influenced by mispricing. These insights highlight the key role of institutional ownership in shaping factor premiums and provide insights for the dynamics of factor pricing.

References

Adrian, Tobias, Erkko Etula, and Tyler Muir, 2014, Financial intermediaries and the cross-section of asset returns, *Journal of Finance* 69, 2557–2596.

Akbas, Ferhat, Will J Armstrong, Sorin Sorescu, and Avanidhar Subrahmanyam, 2015, Smart money, dumb money, and capital market anomalies, *Journal of Financial Economics* 118, 355–382.

Antoniou, Constantinos, John A. Doukas, and Avanidhar Subrahmanyam, 2016, Investor sentiment, beta, and the cost of equity capital, *Management Science* 62, 347–367.

Artiga González, Tanja, Iman van Lelyveld, and Katarína Lučivjanská, 2020, Pension fund equity performance: Patience, activity or both?, *Journal of Banking & Finance* 115, 105812.

Asness, Clifford S, Andrea Frazzini, and Lasse Heje Pedersen, 2019, Quality minus junk, *Review of Accounting Studies* 24, 34–112.

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

Barber, Brad M., and Terrance Odean, 2008, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies* 21, 785–818.

Basak, Suleyman, and Anna Pavlova, 2013, Asset prices and institutional investors, *American Economic Review* 103, 1728–1758.

Belo, Frederico, 2010, Production-based measures of risk for asset pricing, *Journal of Monetary Economics* 57, 146–163.

Belo, Frederico, Jun Li, Xiaoji Lin, and Xiaofei Zhao, 2017, Labor-force heterogeneity and asset prices: The importance of skilled labor, *Review of Financial Studies* 30, 3669–3709.

Ben-Rephael, Azi, Zhi Da, and Ryan D. Israelsen, 2017, It depends on where you search: Institutional investor attention and underreaction to news, *Review of Financial Studies* 30, 3009–3047.

Black, Fischer, 1972, Capital market equilibrium with restricted borrowing, *Journal of Business* 45, 444–455.

Boehmer, Ekkehart, and Eric K. Kelley, 2009, Institutional investors and the informational efficiency of prices, *Review of Financial Studies* 22, 3563–3594.

Breeden, Douglas T, 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics* 7, 265–296.

Calluzzo, Paul, Fabio Moneta, and Selim Topaloglu, 2019, When anomalies are publicized broadly, do institutions trade accordingly?, *Management Science* 65, 4555–4574.

Cao, Jie, Bing Han, and Qinghai Wang, 2017, Institutional investment constraints and stock prices, *Journal of Financial and Quantitative Analysis* 52, 465–489.

Chen, Huafeng Jason, Hernán Ortiz-Molina, and Siliang Stacy Zhang, 2011, Average stock variance and market returns: Evidence of time-varying predictability at the daily frequency, *Journal of Portfolio Management* 37, 86–95.

Chen, Nai-Fu, Richard Roll, and Stephen A Ross, 1986, Economic forces and the stock market, *Journal of Business* 383–403.

Chen, Zhuo, Bibo Liu, Huijun Wang, Zhengwei Wang, and Jianfeng Yu, 2020a, Characteristics-based factors, *Available at SSRN 3112835* .

Chen, Zhuo, Bibo Liu, Huijun Wang, Zhengwei Wang, and Jianfeng Yu, 2020b, Investor sentiment and the pricing of characteristics-based factors, *Available at SSRN* .

Coggin, T. Daniel, Frank J. Fabozzi, and Shafiqur Rahman, 1993, The investment performance of u.s. equity pension fund managers: An empirical investigation, *Journal of Finance* 48, 1039–1055.

Cohen, Randolph B, Christopher Polk, and Tuomo Vuolteenaho, 2005, Money illusion in the stock market: The modigliani-cohn hypothesis, *Quarterly Journal of Economics* 120, 639–668.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011, In search of attention, *Journal of Finance* 66, 1461–1499.

Daniel, Kent, David Hirshleifer, and Lin Sun, 2020a, Short-and long-horizon behavioral factors, *Review of Financial Studies* 33, 1673–1736.

Daniel, Kent, Lira Mota, Simon Rottke, and Tano Santos, 2020b, The cross-section of risk and returns, *Review of Financial Studies* 33, 1927–1979.

Engelberg, Joseph, R David McLean, and Jeffrey Pontiff, 2018, Anomalies and news, *Journal of Finance* 73, 1971–2001.

Fama, Eugene F, and Kenneth R French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

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

Fama, Eugene F, and Kenneth R French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fama, Eugene F, and Michael R Gibbons, 1984, A comparison of inflation forecasts, *Journal of Monetary Economics* 13, 327–348.

Ferreira, Miguel A., and Pedro Matos, 2008, The colors of investors' money: The role of institutional investors around the world, *Journal of Financial Economics* 88, 499–533, Darden - JFE Conference Volume: Capital Raising in Emerging Economies.

Frazzini, Andrea, and Lasse Heje Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1–25.

French, Kenneth R, G William Schwert, and Robert F Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 19, 3–29.

Gilchrist, Simon, and Egon Zakrajšek, 2012, Credit spreads and business cycle fluctuations, *American Economic Review* 102, 1692–1720.

Gompers, Paul A, and Andrew Metrick, 2001, Institutional investors and equity prices, *Quarterly Journal of Economics* 116, 229–259.

Gonçalves, Andrei S, Chen Xue, and Lu Zhang, 2020, Aggregation, capital heterogeneity, and the investment capm, *Review of Financial Studies* 33, 2728–2771.

Grinblatt, Mark, and Matti Keloharju, 2009, Sensation seeking, overconfidence, and trading activity, *Journal of Finance* 64, 549–578.

Guercio, Diane Del, and Jennifer Hawkins, 1999, The motivation and impact of pension fund activism, *Journal of Financial Economics* 52, 293–340.

Harvey, Campbell R, 1986, *Recovering Expectations of Consumption Growth from an Equilibrium Model of the Term Structure of Interest Rates*, Ph.D. thesis, University of Chicago, Graduate School of Business.

Hendershott, Terrence, Dmitry Livdan, and Dominik Rösch, 2020, Asset pricing: A tale of night and day, *Journal of Financial Economics* 138, 635–662.

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

Hou, Kewei, Haitao Mo, Chen Xue, and Lu Zhang, 2021, An augmented q-factor model with expected growth, *Review of Finance* 25, 1–41.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650–705.

Huang, Shiyang, Yang Song, and Hong Xiang, forthcoming, Noise trading and asset pricing factors, *Management Science* .

Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional capm and the cross-section of expected returns, *Journal of Finance* 51, 3–53.

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

Jermann, Urban J, 1998, Asset pricing in production economies, *Journal of Monetary Economics* 41, 257–275.

Ke, Bin, and Santhosh Ramalingegowda, 2005, Do institutional investors exploit the post-earnings announcement drift?, *Journal of Accounting and Economics* 39, 25–53.

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

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?, *Journal of Finance* 64, 1889–1933.

Lewellen, Jonathan, 2011, Institutional investors and the limits of arbitrage, *Journal of Financial Economics* 102, 62 – 80.

Locke, Peter R., and Steven C. Mann, 2005, Professional trader discipline and trade disposition, *Journal of Financial Economics* 76, 401–444.

Long, J. Bradford De, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703–738.

Lou, Dong, 2012, A flow-based explanation for return predictability, *Review of Financial Studies* 25, 3457–3489.

Lucas, Robert E, 1978, Asset prices in an exchange economy, *Econometrica: Journal of the Econometric Society* 1429–1445.

McLean, R David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *Journal of Finance* 71, 5–32.

McLean, R David, Jeffrey Pontiff, and Christopher Reilly, 2020, Taking sides on return predictability, *Georgetown McDonough School of Business Research Paper* .

Mehra, Rajnish, and Edward C Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145–161.

Merton, Robert C, et al., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.

Muravyev, Dmitriy, Neil D. Pearson, and Joshua M. Pollet, 2022, Is there a risk premium in the stock lending market? evidence from equity options, *Journal of Finance* 77, 1787–1828.

Nagel, Stefan, 2005, Short sales, institutional investors and the cross-section of stock returns, *Journal of Financial Economics* 78, 277–309.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Perez-Quiros, Gabriel, and Allan Timmermann, 2000, Firm size and cyclical variations in stock returns, *Journal of Finance* 55, 1229–1262.

Sadka, Ronnie, 2006, Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk, *Journal of Financial Economics* 80, 309–349.

Santos, Tano, and Pietro Veronesi, 2005, Labor income and predictable stock returns, *Review of Financial Studies* 19, 1–44.

Savor, Pavel, and Mungo Wilson, 2014, Asset pricing: A tale of two days, *Journal of Financial Economics* 113, 171–201.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Shen, Junyan, Jianfeng Yu, and Shen Zhao, 2017, Investor sentiment and economic forces, *Journal of Monetary Economics* 86, 1–21.

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

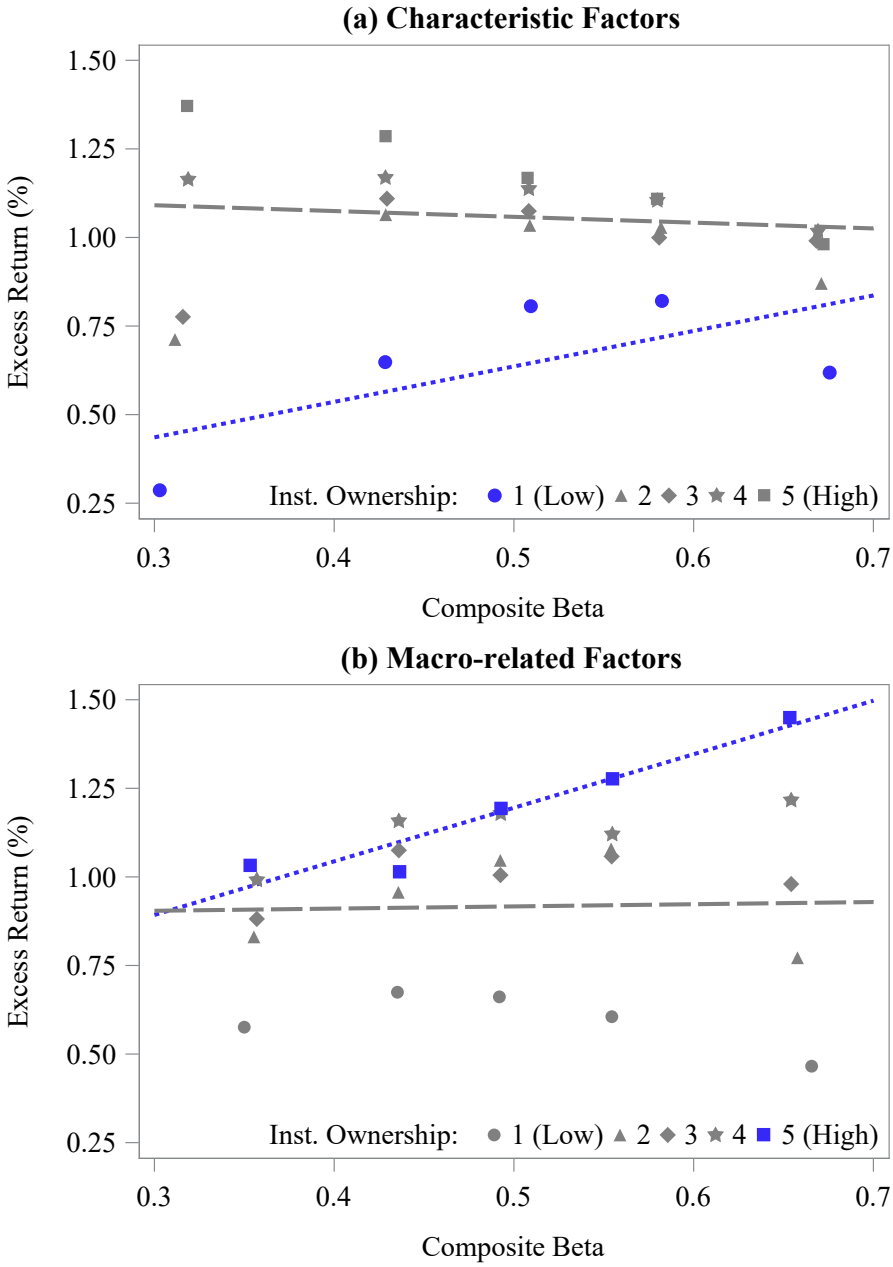
Stambaugh, Robert F, Jianfeng Yu, and Yu Yuan, 2014, The long of it: Odds that investor sentiment spuriously predicts anomaly returns, *Journal of Financial Economics* 114, 613–619.

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

Stambaugh, Robert F, and Yu Yuan, 2017, Mispricing factors, *Review of Financial Studies* 30, 1270–1315.

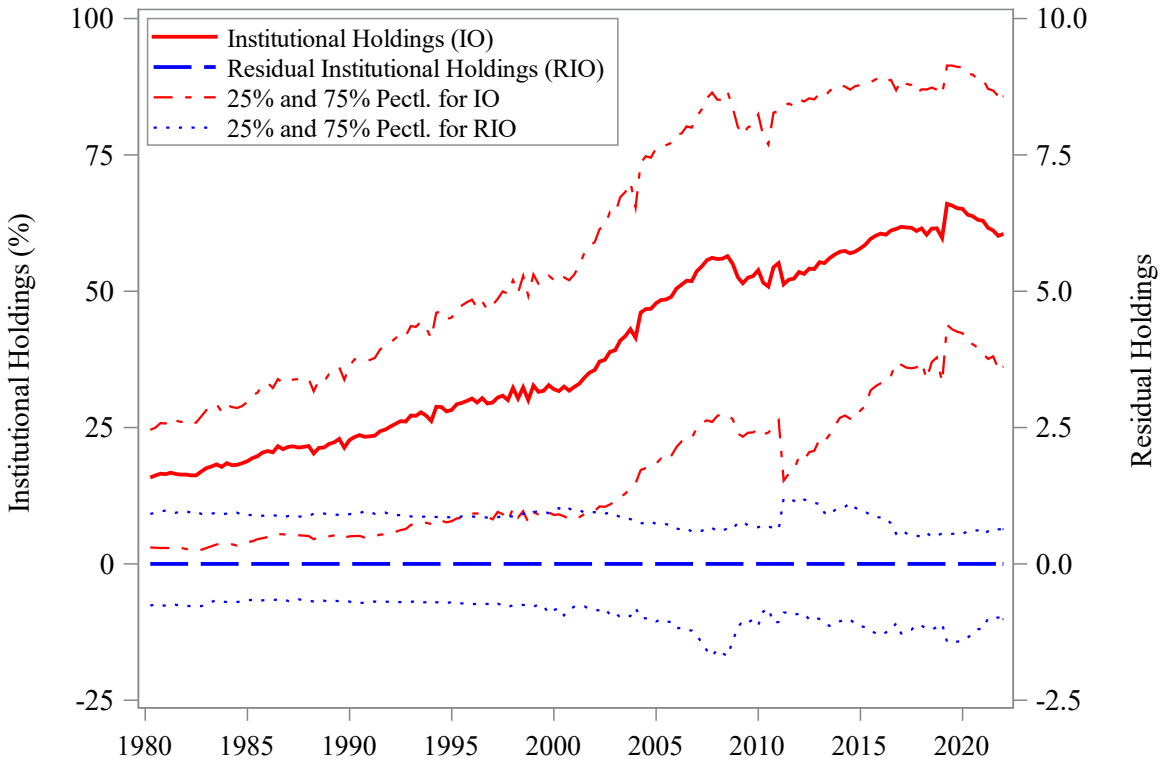
Yu, Jialin, 2011, Disagreement and return predictability of stock portfolios, *Journal of Financial Economics* 99, 162–183.

Figure 1: **Beta-Return Relations of Stocks with Different Institutional Ownership**



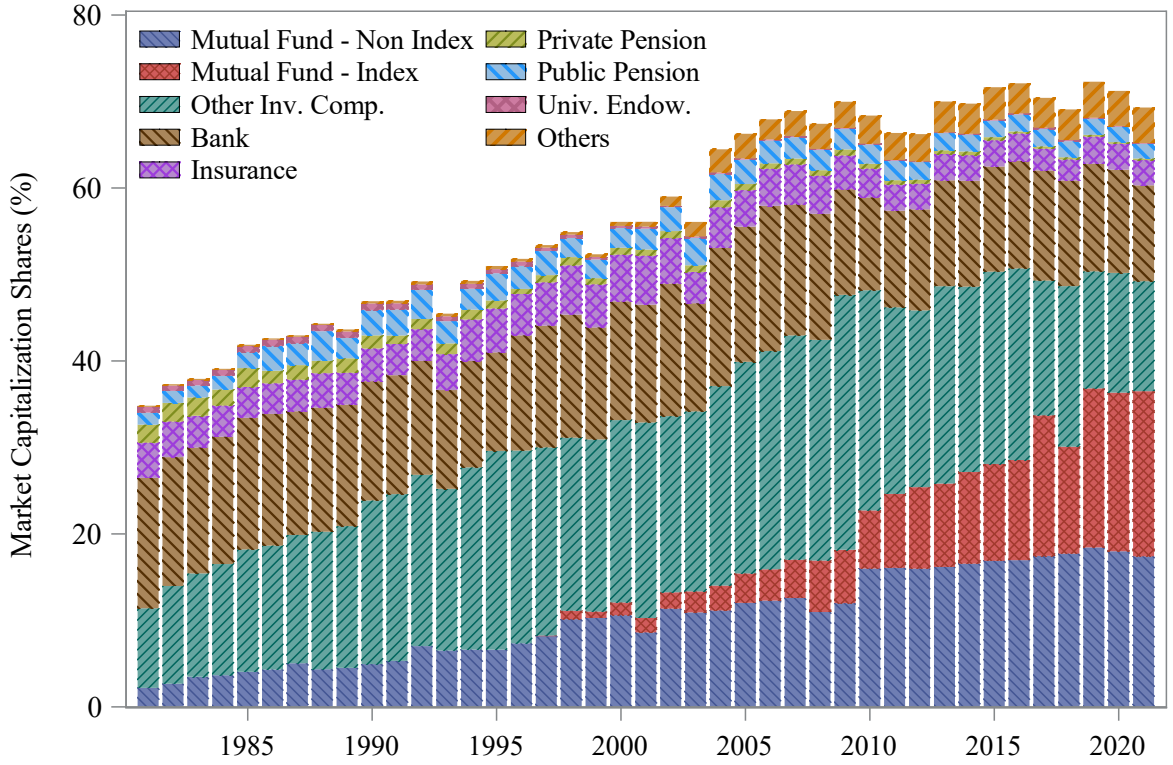
This figure shows the beta-return relationships for portfolios with varying levels of institutional ownership. At the end of each month t, stocks are independently double-sorted into 5×5 groups based on their residual institutional ownership levels $\text{RIO}_{i,t-1}$ estimated from regression (2), and composite factor exposures $\beta^{comp}_{i,j,t-1}$ estimated from regression (4). For both sorting variables, only values up to $t-1$ are used to ensure data availability at the portfolio formation month t. Figure (a) plots the average excess returns in month $t+1$ and the average composite beta $\beta^{comp}_{i,j,t-1}$ for each portfolio sorted by characteristic factor composite beta, while Figure (b) shows the same for portfolios sorted by macro-related factor composite beta. Following Nagel (2005), equal-weighted portfolio returns are reported to mitigate the size effect. The sample period spans from January 1980 to December 2021.

Figure 2: **Time Seires of Institutional Holdings**



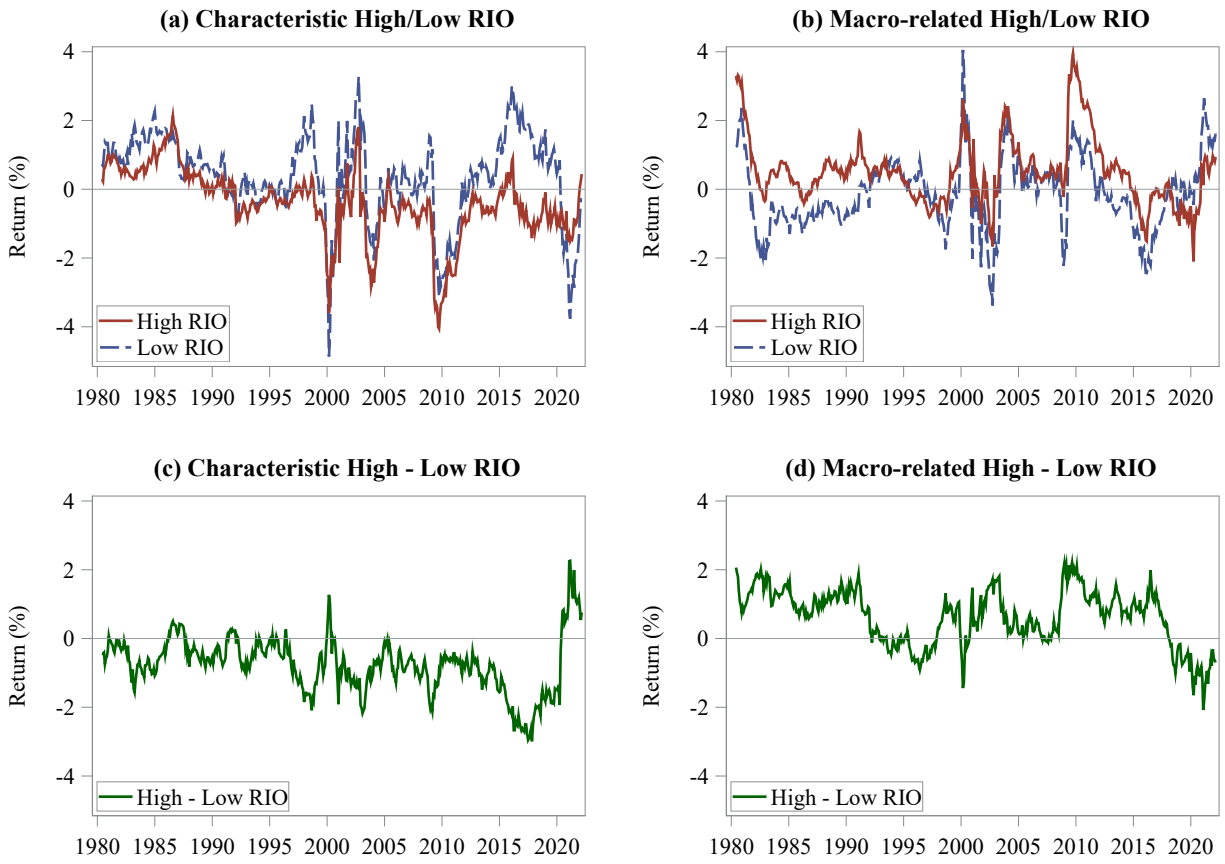
This figure shows the time series of cross-sectional mean and first and third quartiles of institutional ownership (IO, left axis) and residual institutional ownership (RIO, right axis). RIO is residual institutional holdings after exempting the effect of firm size on IO, as estimated following equation (2). The sample period spans from 1980Q1 to 2021Q4.

Figure 3: **Institutional Holdings Decomposition**



This figure shows the annual market capitalization shares of institutional types. Institutions are categorized into into seven types: investment companies and advisors (further divided into non-index mutual funds, index mutual funds, and other investment companies), banks, insurance companies, public pension funds, private (coporate) pension funds, university and foundation endowments, and others. The sample period spans from 1981 to 2021.

Figure 4: **Time Series of Factor Premium Differences**



This figure shows the the time series of factor premiums estimated for stocks with high and low residual institutional ownership (RIO), along with their differences. At the end of every month t, stocks are independently double-sorted into 5×5 groups based on their residual institutional ownership levels $\text{RIO}_{i,t-1}$ estimated from regression (2) and pre-formation factor exposure $\beta^{pre}_{i,k,t-1}$ for each factor k estimated from regression (3). For both sorting variables, we only use values up to $t-1$ to ensure data availability at the portfolio formation month t. Following Nagel (2005), we report equal-weighted portfolio returns to mitigate the size effect. Figure (a) and (b) shows the factor premiums for characteristic and macro-related factors, measured as the excess returns of high-minus-low portfolios constructed by taking a long position in the highest β portfolios and a short position in the lowest β portfolios, in high (solid lines) and low (dashed lines) RIO stocks. Figure (c) and (d) show the factor premium differences between high and low RIO stocks. To minimize noise in short-term returns, we employ an exponential weighted moving average (EWMA) model with a decay parameter of 0.95. The sample period spans from January 1980 to December 2021.

Table 1: **Summary Statistics of Institutional Holdings**

Panel A. Institutional ownership (IO)

Decade	Mean	Std	P10	P50	P90	# Firms	Coverage
1980s	0.20	0.19	0.01	0.14	0.48	4,492	0.80
1990s	0.29	0.24	0.02	0.23	0.65	6,101	0.93
2000s	0.45	0.31	0.04	0.43	0.89	4,918	0.98
2010s	0.58	0.33	0.06	0.66	0.96	3,552	0.96
2020-2021	0.62	0.30	0.13	0.70	0.96	3,682	0.95
Full Sample	0.37	0.31	0.02	0.30	0.85	4,714	0.92

Panel B. Residual institutional ownership (RIO)

Decade	Mean	Std	P10	P50	P90	# Firms	Coverage
1980s	0.00	1.43	-1.84	0.20	1.52	4,492	0.80
1990s	0.00	1.47	-1.76	0.15	1.54	6,101	0.93
2000s	0.00	2.08	-2.07	-0.13	1.80	4,918	0.98
2010s	0.00	2.62	-2.33	-0.18	2.26	3,552	0.96
2020-2021	0.00	2.12	-1.97	-0.26	1.72	3,682	0.95
Full Sample	0.00	1.90	-1.97	0.03	1.66	4,714	0.92

This table shows the summary statistics of the institutional ownership (IO) and residual institutional ownership (RIO, calculated based on equation (2)) for each decade, as well as for the full sample period, from January 1980 to December 2021. # Firms is the time series average of the quarterly number of common stocks in the CRSP database with valid ownership data reported in the 13F. Coverage is the ratio of #Firms and the total number of common stocks in the CRSP database.

Table 2: **Factor Descriptions**

Panel A. Characteristic factors

	Abbr.	Description	Literature
1	SMB	Fama-French size factor	Fama and French (1993)
2	HML	Fama-French book-to-market factor	Fama and French (1993)
3	MOM	Momentum	Jegadeesh and Titman (1993)
4	CMA	Fama-French investment factor	Fama and French (2015)
5	RMW	Fama-French profitability factor	Fama and French (2015)
6	R_{IA}	q5 investment factor	Hou et al. (2015)
7	R_{ROE}	q5 ROE factor	Hou et al. (2015)
8	R_{EG}	q5 expected investment growth factor	Hou et al. (2021)
9	PMU	Gross profitability factor	Novy-Marx (2013)
10	MGMT	Mispricing factor - management	Stambaugh and Yuan (2017)
11	PERF	Mispricing factor - performance	Stambaugh and Yuan (2017)
12	QMJ	Quality factor	Asness et al. (2019)

Panel B. Macro-related factors

	Abbr.	Description	Literature
1	CON	Consumption growth	Lucas (1978)
2	TFP	Total factor productivity growth	Jermann (1998)
3	IPG	Industrial production growth	Chen et al. (1986)
4	DEF	Default Spread (BAA - AAA) change	Gilchrist and Zakrajšek (2012)
5	TERM	Term Spread (T10Y-T1Y) change	Harvey (1986)
6	UI	Unexpected inflation	Fama and Gibbons (1984)
7	DEI	Change of expected inflation	Fama and Gibbons (1984)
8	MKT	Market excess return	Sharpe (1964)
9	VOL	Aggregate market volatility	French et al. (1987)
10	LAB	Labor income growth	Santos and Veronesi (2005)

This table shows an overview of the 22 factors studied in the paper. Factors are divided into two groups: characteristic factors in Panel A, and macro-related factors in Panel B. This table includes the factors' abbreviations, descriptions, and related literature. Appendix A provides details of factor data sources and formation process.

Table 3: **Failure of Factor Pricing**

Panel A. Excess returns of beta-sorted portfolios

(a) ***Characteristic***	SMB	HML	MOM	CMA	RMW	R_{IA}	R_{ROE}	R_{EG}	PMU	MGMT	PERF	QMJ	COMP
High β	1.01	1.17	0.66	0.96	0.75	1.00	0.63	0.92	1.03	0.92	0.75	0.78	0.81
	[2.08]	[3.25]	[1.69]	[2.71]	[2.27]	[2.73]	[2.02]	[3.13]	[2.67]	[3.35]	[2.09]	[3.31]	[3.28]
Low β	0.88	0.83	1.10	0.85	0.92	0.82	1.08	0.85	0.62	0.75	1.01	0.80	0.85
	[3.89]	[1.79]	[2.51]	[1.84]	[1.97]	[1.78]	[2.22]	[1.71]	[1.38]	[1.55]	[2.25]	[1.74]	[1.86]
High-Low	0.13	0.35	-0.44*	0.10	-0.18	0.18	-0.45	0.07	0.41	0.17	-0.26	-0.11	-0.05
	[0.40]	[1.06]	[-1.84]	[0.34]	[-0.58]	[0.61]	[-1.37]	[0.22]	[1.61]	[0.47]	[-0.90]	[-0.30]	-0.05

(b) ***Macro-related***	CON	TFP	IPG	DEF	TERM	UI	DEI	MKT	VOL	LAB			COMP
High β	1.03	1.02	1.15	1.06	0.93	0.92	1.02	1.04	1.03	0.91			0.95
	[2.28]	[2.18]	[2.50]	[2.29]	[2.34]	[2.04]	[2.40]	[2.06]	[2.20]	[2.07]			[2.11]
Low β	1.06	0.88	0.84	0.92	0.95	0.95	0.92	0.77	0.77	1.03			0.79
	[2.75]	[2.39]	[2.36]	[2.59]	[2.33]	[2.53]	[2.28]	[3.06]	[2.45]	[2.56]			[3.21]
High-Low	-0.04	0.14	0.31	0.14	-0.01	-0.02	0.10	0.27	0.27	-0.12			0.16
	[-0.24]	[0.69]	[1.53]	[0.70]	[-0.08]	[-0.11]	[0.74]	[0.80]	[1.10]	[-0.69]			[0.51]

Panel B. Post-formation full sample beta of beta-sorted portfolios

(a) ***Characteristic***	SMB	HML	MOM	CMA	RMW	R_{IA}	R_{ROE}	R_{EG}	PMU	MGMT	PERF	QMJ
High β	2.10	0.39	-0.36	-0.26	-0.51	-0.13	-0.57	-1.02	0.02	-0.36	-0.48	-0.97
	[14.22]	[2.15]	[-1.45]	[-1.41]	[-3.54]	[-0.62]	[-4.91]	[-6.73]	[0.12]	[-4.12]	[-3.92]	[-7.60]
Low β	0.61	-1.19	-0.65	-1.85	-2.66	-1.99	-2.41	-2.90	-1.10	-2.03	-1.28	-3.35
	[5.13]	[-4.02]	[-3.85]	[-4.72]	[-13.34]	[-4.93]	[-10.47]	[-11.60]	[-4.89]	[-15.60]	[-9.12]	[-19.27]
High-Low	1.49***	1.58***	0.29***	1.60***	2.15***	1.86***	1.84***	1.88***	1.12***	1.67***	0.80***	2.38***
	[10.34]	[9.07]	[2.01]	[4.90]	[12.01]	[5.91]	[8.59]	[7.84]	[8.93]	[12.14]	[9.24]	[11.29]

(b) ***Macro-related***	CON	TFP	IPG	DEF	TERM	UI	DEI	MKT	VOL	LAB	SUR
High β	0.11	0.17	1.44	13.92	3.47	1.95	-1.67	1.65	0.38	-0.29	0.01
	[0.20]	[0.70]	[4.82]	[3.68]	[2.14]	[1.40]	[-2.15]	[16.69]	[6.14]	[-0.57]	[0.24]
Low β	-0.04	0.17	1.39	9.98	2.13	0.03	-1.45	0.69	0.25	-1.06	0.00
	[-0.13]	[0.91]	[3.96]	[3.08]	[1.40]	[0.04]	[-1.91]	[12.55]	[-6.40]	[-1.44]	[0.11]
High-Low	0.15	0.00	0.04	3.94***	1.34***	1.92***	-0.22	0.95***	0.13***	0.77*	0.01
	[0.57]	[0.00]	[0.22]	[2.85]	[2.58]	[3.13]	[-0.88]	[10.50]	[3.58]	[1.70]	[0.39]

This table presents the monthly excess returns and post-formation betas for the highest beta, lowest beta, and the high-minus-low long-short portfolios constructed based on factor loadings of 12 characteristic factors and 10 macro-related factors. At the end of each month t, stocks are sorted into deciles based on their pre-formation beta ($\beta^{pre}_{i,k,t-1}$) as estimated in regression (3) for each of the 22 factors. To ensure data availability, only information up to month $t-1$ is used for portfolio formation at time t. High-minus-low portfolios are then constructed by going long on the highest beta deciles (High β and shorting the lowest beta deciles (Low β). Portfolio returns are equal-weighted. Panel A shows the average excess returns of formed portfolios. Panel B reports the post-formation beta ($\beta^{post}_{p,k}$) of the formed portfolios following regression (5). For the two counter-cyclical macro-related factors, default spreads (DEF) and market volatility (VOL), the sign of β is reversed. The sample period is from January 1980 to December 2021. The reported t-statistics use Newey-West standard errors, adjusting for serial correlations.

Table 4: **Factor Premiums in Stocks with High and Low RIO**

Panel A. Characteristic factors						
	Low RIO	2	3	4	High RIO	High - Low
High β	0.63	0.88	1.01	1.03	0.99	0.36***
	[2.90]	[4.14]	[4.30]	[3.78]	[3.44]	[2.42]
Low β	0.32	0.75	0.80	1.19	1.40	1.09***
	[0.68]	[1.77]	[1.98]	[2.88]	[3.27]	[6.18]
High - Low	0.32	0.13	0.20	-0.16	-0.41	-0.72***
	[0.84]	[0.39]	[0.73]	[-0.66]	[-1.52]	[-3.23]
CAPM α	0.86	0.63	0.62	0.25	-0.04	-0.90***
	[2.52]	[2.00]	[2.39]	[1.06]	[-0.15]	[-4.26]
Panel B. Macro-related factors						
	Low RIO	2	3	4	High RIO	High - Low
High β	0.49	0.79	0.99	1.23	1.47	0.98***
	[1.08]	[1.96]	[2.52]	[2.98]	[3.35]	[5.49]
Low β	0.61	0.85	0.91	1.01	1.06	0.45***
	[2.73]	[3.81]	[3.85]	[3.94]	[3.54]	[3.07]
High - Low	-0.12	-0.06	0.08	0.22	0.41	0.54***
	[-0.36]	[-0.22]	[0.33]	[0.87]	[1.63]	[2.63]
CAPM α	-0.61	-0.53	-0.27	-0.14	0.02	0.63***
	[-1.96]	[-1.97]	[-1.17]	[-0.61]	[0.07]	[2.96]

This table shows the excess returns and CAPM α of double-sorted portfolios based on composite beta and residual institutional ownership (RIO). At the end of every month t, stocks are independently double-sorted into 5×5 groups based on their residual institutional ownership levels $\text{RIO}_{i,t-1}$ estimated from regression (2) and composite factor exposures $\beta^{comp}_{i,j,t-1}$ estimated from regression (4). Panel A reports the results of portfolios sorted by characteristic factor composite betas, while Panel B focuses on portfolios sorted by macro-related factor composite betas. For both sorting variables, we only use values up to $t-1$ to ensure data availability at the portfolio formation month t. Following Nagel (2005), we report equal-weighted portfolio returns to mitigate the size effect. The table shows excess returns of 10 portfolios with highest and lowest β^{comp} for each of the five RIO quintitles. High-Low rows represent excess returns for portfolios constructed by taking a long position in the highest β^{comp} portfolios and a short position in the lowest β^{comp} portfolios within each RIO quintitles. High-Low columns show return and CAPM α differences between highest and lowest RIO portfolios. The CAPM α is estimated using the regression: $r_{p,t} - r_{f,t} = \alpha_p + \beta_p \times (r_{mkt,t} - r_{f,t}) + \epsilon_{p,t}$ for high-minus-low portfolio p. The sample period spans from January 1980 to December 2021. The reported t-statistics use Newey-West standard errors, adjusting for serial correlations.

Table 5: **Single Factor Premiums in Stocks with High and Low RIO**

		(A) Characteristic factors				*(B) Macro-related factors*		
		High - Low β				High - Low β		
		High RIO	Low RIO	High - Low		High RIO	Low RIO	High - Low
AVG 1-12		-0.26	0.09	-0.36***	AVG 1-10	0.22	-0.05	0.27***
		[-1.77]	[0.47]	[-3.06]		[1.80]	[-0.37]	[3.45]
AVG 2-12		-0.30	0.13	-0.43***	AVG 1-6	0.26	-0.06	0.32***
		[-1.71]	[0.51]	[-3.11]		[1.78]	[-0.33]	[3.18]
AVG 10-12		-0.42	0.18	-0.59***	AVG 1-8	0.24	-0.06	0.30***
		[-1.84]	[0.62]	[-3.60]		[1.89]	[-0.41]	[3.47]
1	SMB	0.15	-0.25	0.40*	CON	-0.07	-0.14	0.07
		[0.65]	[-0.74]	[1.90]		[-0.51]	[-0.94]	[0.43]
2	HML	0.17	0.43	-0.26	TFP	0.43	-0.09	0.53***
		[0.57]	[1.30]	[-0.99]		[2.04]	[-0.48]	[3.00]
3	MOM	-0.41	-0.41	0.00	IPG	0.42	0.20	0.22
		[-1.86]	[-1.73]	[0.00]		[2.21]	[0.99]	[1.18]
4	CMA	-0.10	0.12	-0.22	MKT	0.33	-0.16	0.49**
		[-0.38]	[0.37]	[-0.86]		[1.17]	[-0.48]	[2.47]
5	RMW	-0.60	0.06	-0.66***	VOL	0.43	0.08	0.35
		[-2.38]	[0.17]	[-2.99]		[2.06]	[0.28]	[1.62]
6	R_{IA}	-0.09	0.15	-0.24	LAB	0.04	-0.22	0.26
		[-0.34]	[0.47]	[-0.89]		[0.21]	[-1.34]	[1.54]
7	R_{ROE}	-0.69	-0.11	-0.58***	DEF	0.24	-0.20	0.44**
		[-2.60]	[-0.37]	[-3.10]		[1.24]	[-0.86]	[2.45]
8	R_{EG}	-0.27	0.30	-0.57***	TERM	0.07	0.03	0.04
		[-0.95]	[0.94]	[-2.97]		[0.44]	[0.17]	[0.22]
9	PMU	-0.10	0.31	-0.41**	UI	0.20	-0.12	0.32*
		[-0.43]	[1.16]	[-2.21]		[0.98]	[-0.58]	[1.92]
10	MGMT	-0.08	0.36	-0.44	DEI	0.06	0.09	-0.02
		[-0.27]	[0.93]	[-1.61]		[0.44]	[0.58]	[-0.17]
11	PERF	-0.57	-0.06	-0.51**				
		[-2.32]	[-0.23]	[-2.43]				
12	QMJ	-0.60	0.23	-0.83***				
		[-2.01]	[0.63]	[-4.02]				

This table shows the excess returns of double-sorted portfolios based on single factor betas and residual institutional ownership. Panel A reports results for characteristic factors, while Panel B focuses on macro-related factors. At the end of every month t, stocks are independently double-sorted into 5×5 groups based on their residual institutional ownership levels $\text{RIO}_{i,t-1}$ estimated from regression (2) and pre-formation factor exposure $\beta^{pre}_{i,k,t-1}$ for each factor k estimated from regression (3). For both sorting variables, we only use values up to $t-1$ to ensure data availability at the portfolio formation month t. Following Nagel (2005), we report equal-weighted portfolio returns to mitigate the size effect. The table shows the excess returns of high-minus-low portfolios, constructed by taking a long position in the highest β portfolios and a short position in the lowest β portfolios, within the highest and lowest RIO quintiles as well as their differences. The AVG portfolios represent average portfolios based on single-factor portfolios, while AVG i-j refers to the averages from the ith to jth factor-based portfolios as numbered in the table. The sample period is from January 1980 to December 2021. The reported t-statistics use Newey-West standard errors, adjusting for serial correlations.

Table 6: **Fama-Macbeth Regressions of RIO Impacts on Beta-Return Relations**

		(A) *Characteristic factors*				(B) *Macro-related factors*		
		$\beta \times RIO$	β	RIO		$\beta \times RIO$	β	RIO
	COMP	-0.11***	0.01	0.22	COMP	0.11***	0.06	0.23
		[-4.02]	[0.07]	[6.54]		[5.08]	[0.60]	[6.72]
1	SMB	0.08***	0.03	0.23	CON	0.05***	-0.01	0.25
		[3.75]	[0.31]	[6.81]		[2.94]	[-0.37]	[6.85]
2	HML	-0.06**	0.13	0.24	TFP	0.07***	0.01	0.25
		[-2.21]	[1.19]	[6.96]		[3.98]	[0.23]	[6.79]
3	MOM	-0.07***	-0.09	0.24	IPG	0.13***	0.05	0.23
		[-3.00]	[-1.20]	[6.64]		[6.40]	[0.73]	[6.48]
4	CMA	-0.05	0.04	0.24	DEF	0.14***	0.05	0.23
		[-1.58]	[0.38]	[6.89]		[6.32]	[0.62]	[6.78]
5	RMW	-0.10***	-0.07	0.24	TERM	0.07***	0.02	0.25
		[-3.55]	[-0.69]	[6.95]		[3.22]	[0.25]	[6.87]
6	R_{IA}	-0.04	0.06	0.24	UI	0.08***	0.02	0.24
		[-1.21]	[0.59]	[6.83]		[3.87]	[0.25]	[6.95]
7	R_{ROE}	-0.09***	-0.12	0.23	DEI	-0.11***	-0.02	0.24
		[-3.72]	[-1.15]	[6.88]		[-5.16]	[-0.26]	[6.69]
8	R_{EG}	-0.10***	0.03	0.23	MKT	0.06**	0.03	0.24
		[-4.24]	[0.25]	[6.95]		[2.52]	[0.31]	[6.80]
9	PMU	-0.07***	0.04	0.23	VOL	0.09***	0.07	0.23
		[-3.52]	[0.54]	[6.56]		[3.97]	[0.77]	[6.67]
10	MGMT	-0.09***	0.06	0.23	LAB	-0.08***	-0.04	0.25
		[-3.00]	[0.51]	[6.81]		[-3.71]	[-0.80]	[6.72]
11	PERF	-0.07***	-0.07	0.24				
		[-3.16]	[-0.74]	[6.69]				
12	QMJ	-0.12***	-0.03	0.22				
		[-5.02]	[-0.22]	[6.81]				

This table shows the results of Fama-MacBeth regressions based on equation (6), reporting the time-series averages of the interaction coefficients b_1 ($\beta \times$ RIO), b_2 (β), and b_3 (RIO) for each factor. Residual institutional ownership, $\text{RIO}_{i,t}$, for stock i is estimated using regression (2). The post-formation beta, $\beta^{post}_{i,k,t}$, for stock i on factor k in month t, is defined as the post-formation beta, $\beta^{post}_{p,k}$, of the pre-formation beta-sorted portfolio p (as estimated using equation (5)) to which stock i belongs in month t. The composite post-formation betas are calculated as the average percentile rankings of single-factor post-formation betas. Both RIO and post-formation betas are standardized to have a cross-sectional mean of 0 and a standard deviation of 1 prior to forming the interaction terms. The sample period spans from January 1980 to December 2021. The reported t-statistics use Newey-West standard errors, adjusting for serial correlations.

Table 7: **Factor Premiums in Stocks Held by Different Types of Institutions**

Type	Mkt Cap	(A) *Characteristic*			(B) *Macro-related*		
		High - Low β			High - Low β		
		High RIO	Low RIO	High - Low	High RIO	Low RIO	High - Low
All	69.6%	-0.45*	0.30	-0.75***	0.39	-0.18	0.57***
		[-1.68]	[0.78]	[-3.30]	[1.52]	[-0.52]	[2.77]
- Inv. Comp.	49.2%	-0.47*	0.26	-0.74***	0.39	-0.21	0.60***
		[-1.77]	[0.66]	[-3.01]	[1.62]	[-0.62]	[2.84]
- Bank	11.1%	-0.32	0.00	-0.32*	0.3	0.01	0.29
		[-1.07]	[-0.01]	[-1.91]	[1.10]	[0.05]	[1.61]
- Insurance	2.9%	-0.05	0.01	-0.06	0.00	0.04	-0.04
		[-0.19]	[0.03]	[-0.38]	[0.01]	[0.14]	[-0.23]
- Public Pension	1.7%	-0.30	0.24	-0.53***	0.21	-0.19	0.40**
		[-0.97]	[0.71]	[-3.07]	[0.81]	[-0.62]	[2.21]
- Private Pension	0.2%	-0.12	0.02	-0.15	0.06	0.08	-0.02
		[-0.42]	[0.08]	[-0.94]	[0.24]	[0.26]	[-0.11]
- Univ. Endow.	0.1%	-0.04	0.08	-0.12	0.03	-0.21	0.23
		[-0.13]	[0.27]	[-0.72]	[0.11]	[-0.67]	[1.14]
- Others	4.5%	0.00	-0.29	0.28	0.18	0.29	-0.11
		[-0.01]	[-0.85]	[1.35]	[0.64]	[0.93]	[-0.54]
Within Inv. Comp.							
- Mutual Fund (MF)	36.5%	-0.16	0.20	-0.36*	0.20	-0.12	0.32*
		[-0.63]	[0.54]	[-1.72]	[0.82]	[-0.40]	[1.65]
- MF - Index	19.1%	-0.57	-0.39	-0.18	0.34	0.14	0.20
		[-1.14]	[-0.78]	[-0.56]	[0.84]	[0.31]	[0.86]
- MF - Non Index	17.4%	-0.14	0.20	-0.35*	0.20	0.03	0.17
		[-0.57]	[0.56]	[-1.75]	[0.81]	[0.09]	[0.88]
- Other Inv. Comp.	12.7%	-0.47*	0.30	-0.77***	0.41*	-0.25	0.66***
		[-1.67]	[0.76]	[-3.26]	[1.65]	[-0.76]	[3.20]

This table shows the factor premiums in stocks with high and low holdings from different types of institutions. Institutions are classified into seven types: investment companies or independent investment advisors (Inv. Comp.), banks, insurance companies, public pension funds, private (copoorate) pension funds, university and foundation endowments (Univ. Endow.), and others. Investment companies and advisors are further divided into three subtypes: index mutual funds, non-index mutual funds, and other investment companies. We calculate residual institutional ownership (RIO) for each investor type using equation (2), replacing the IO variable with holdings specific to each institution type. For each type of the institution, stocks are independently double-sorted into 5×5 groups based on their RIO and composite factor exposures $\beta^{comp}_{i,j,t-1}$ estimated from regression (4) at the end of every month t. For both sorting variables, we only use values up to $t-1$ to ensure data availability at the portfolio formation month t. Following Nagel (2005), we report equal-weighted portfolio returns to mitigate the size effect. The table shows the excess returns of high-minus-low portfolios, constructed by taking a long position in the highest β portfolios and a short position in the lowest β portfolios, within the highest and lowest RIO quintiles as well as their differences. Sample period spans from January 1981 to December 2021. The reported t-statistics use Newey-West standard errors, adjusting for serial correlations.

Table 8: **Compare the Effects of RIO with Time-series Forces on Beta-Return Relations**

	(A) Characteristic					***(B) Macro-related***				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
$\beta \times$ RIO	-0.13***	-0.15***	-0.15***	-0.15***	-0.14***	0.12***	0.13***	0.12***	0.12***	0.11***
	[-4.28]	[-4.55]	[-4.48]	[-4.38]	[-4.38]	[4.15]	[4.32]	[3.97]	[3.99]	[3.69]
$\beta \times$ Sent	0.46**				0.40*	-0.34**				-0.32*
	[2.28]				[1.73]	[-2.03]				[-1.65]
$\beta \times \Delta$Inf		0.19*			0.16		-0.18**			-0.16
		[1.71]			[1.38]		[-1.97]			[-1.56]
$\beta \times$ Leverage			0.42***		0.35***			-0.32***		-0.27***
			[3.73]		[2.89]			[-3.43]		[-2.65]
$\beta \times$ Disagree				-0.05	0.08				0.03	-0.08
				[-0.39]	[0.47]				[0.34]	[-0.62]
β	-0.02	0.00	0.02	-0.02	0.02	0.07	0.06	0.04	0.06	0.03
	[-0.14]	[-0.03]	[0.16]	[-0.13]	[0.14]	[0.69]	[0.59]	[0.36]	[0.61]	[0.26]
RIO	0.23***	0.24***	0.24***	0.24***	0.23***	0.24***	0.24***	0.24***	0.24***	0.24***
	[6.53]	[6.70]	[6.41]	[6.51]	[6.15]	[6.62]	[6.72]	[6.44]	[6.49]	[6.16]
Sent	-0.46				-0.38	-0.46				-0.39
	[-1.51]				[-1.14]	[-1.51]				[-1.14]
ΔInf		-0.62**			-0.54*		-0.62**			-0.54*
		[-2.43]			[-1.94]		[-2.43]			[-1.94]
Leverage			-0.52*		-0.42			-0.53*		-0.42
			[-1.87]		[-1.43]			[-1.88]		[-1.43]
Disagree				0.22	0.00				0.21	-0.01
				[0.87]	[-0.01]				[0.86]	[-0.02]
Intercept	0.98***	0.96***	0.92***	0.97***	0.91***	0.98***	0.96***	0.92***	0.97***	0.91***
	[3.84]	[3.77]	[3.48]	[3.66]	[3.23]	[3.84]	[3.78]	[3.48]	[3.67]	[3.23]
N	2,020,803	2,020,803	1,869,985	1,957,286	1,806,468	2,020,803	2,020,803	1,869,985	1,957,286	1,806,468
R-Square	0.15	0.14	0.17	0.04	0.35	0.12	0.14	0.15	0.04	0.30

This table shows the results of panel regression (7) to compare the effects of residual institutional ownership (RIO) and time-series variables on the beta-return relationship. Panel A reports results for characteristic factors, while Panel B focuses on macro-related factors. The time-series variables include: (1) Sentiment, obtained from Jeffrey Wurgler's website; (2) Change in monthly inflation; (3)Broker-dealer leverage measure, obtained from Tyler Muir's website with data available up to the third quarter of 2017. (4) Disagreement, estimated using the aggregate standard deviations of long-term earnings-per-share forecasts from the IBES database, with data starting from December 1981. $\beta_{i,j,t}^{comp_post}$ and $\text{RIO}_{i,t}$ are standardized to have a mean of 0 and a standard deviation of 1 across all stocks for each month, while $tsVar_t$ is standardized across time to have a mean of 0 and a standard deviation of 1. Sample periods spans from January 1980 to December 2021. The reported t-statistics use two-way clustered standard errors.

Table 9: **Combined Effects of Sentiment and RIO on Factor Premiums**

Panel A. Characteristic factors

	Low RIO	2	3	4	High RIO	High - Low
High Sent	0.98**	0.89*	0.82**	0.58*	0.21	-0.77***
	[2.00]	[1.95]	[2.16]	[1.90]	[0.66]	[-2.62]
Low Sent	-0.35	-0.63	-0.41	-0.92***	-1.03***	-0.68**
	[-0.68]	[-1.46]	[-1.14]	[-2.60]	[-2.58]	[-2.06]
High - Low	1.33**	1.52**	1.23**	1.50***	1.23**	-0.10
	[1.99]	[2.52]	[2.38]	[3.32]	[2.44]	[-0.23]

Panel B. Macro-related factors

	Low RIO	2	3	4	High RIO	High - Low
High Sent	-0.72	-0.66	-0.21	-0.17	-0.06	0.66**
	[-1.56]	[-1.60]	[-0.57]	[-0.52]	[-0.22]	[2.40]
Low Sent	0.48	0.54	0.38	0.61*	0.89**	0.41
	[1.08]	[1.54]	[1.25]	[1.75]	[2.33]	[1.45]
High - Low	-1.19**	-1.20**	-0.60	-0.78	-0.95**	0.24
	[-1.99]	[-2.30]	[-1.28]	[-1.61]	[-2.01]	[0.67]

This table shows the factor premiums within stocks grouped by RIO levels and across different sentiment periods. At the end of each month t, stocks are double-sorted into 5×5 groups based on their residual institutional ownership levels $\text{RIO}_{i,t-1}$ estimated from regression (2) and composite factor exposures $\beta^{comp}_{i,j,t-1}$ estimated from regression (4). For both sorting variables, we only use values up to $t-1$ to ensure data availability at the portfolio formation month t. Factor premiums are calculated as the excess returns of high-minus-low beta-sorted portfolios within each RIO tier. Following Nagel (2005), we report equal-weighted portfolio returns to mitigate the size effect. The sample is split into high and low sentiment periods, determined by the median level of investor sentiment over the full sample. Panel A reports characteristic factor premiums following high and low sentiment periods across different RIO tiers, while Panel B reports macro-related factor premiums. Investor sentiment data is obtained from Jeffrey Wurgler's website. The reported t-statistics use Newey-West standard errors, adjusting for serial correlations.

Table 10: **Measures of Short-Sale Constraints**

Panel A. Summary statistics of short-sale constraint measures

		Short Interest	Utilization	Indicative Fee
	Mean	2.80	15.48	2.50
	Std	5.21	20.63	8.18
	Number of Months	504	236	236
	Number of Firms	2395	3212	3072

Panel B. Correlations between short-sale constraint measures and IO/RIO

		Short Interest	Utilization	Indicative Fee
	Short Interest	1		
	Utilization	0.45	1	
Corr	Indicative Fee	0.11	0.42	1
	IO	0.13	-0.06	-0.31
	RIO	0.13	0.06	-0.08

This table shows the summary statistics for three short-sale constraint measures - short interest, utilization, and indicative fee - and their correlations with IO and RIO, as well as their inter-correlations. Short interest is the ratio of shares held short to total outstanding shares, with data obtained from the Compustat Supplemental Short Interest File. Utilization represents the proportion of a stock's value on loan relative to its total lendable value in the borrowing market. Indicative fee reflects the borrowing cost of equities in the borrowing market. Data for utilization and indicative fees are obtained from Markit Securities Finance Analytics. The short interest data spans January 1980 to December 2021, while utilization and indicative fee data are available from May 2002 onward. "Number of firms" refers to the average number of firms with available short-sale constraint measures in each month.

Table 11: **Factor Premium Differences with Control of Short Sale Constraints**

	(A) *Characteristic*				(B) *Macro-related*			
Panel A: Stocks with high short-sale constraints								
SSC Measures	High RIO	Low RIO	High - Low Ex. Ret	High - Low Alpha	High RIO	Low RIO	High - Low Ex. Ret	High - Low Alpha
(1) *Short Interest*	-0.31	0.10	-0.41**	-0.53***	0.21	0.1	0.11	0.18
	[-1.35]	[0.34]	[-2.49]	[-3.19]	[0.89]	[0.41]	[0.61]	[1.02]
(2) *Utilization*	-0.67*	-0.23	-0.44*	-0.55**	0.42	0.15	0.27	0.29
	[-1.79]	[-0.48]	[-1.76]	[-2.33]	[1.00]	[0.35]	[1.14]	[1.16]
(3) *Indicative Fee*	-0.84**	-0.13	-0.71**	-0.87***	0.42	0.17	0.25	0.28
	[-2.01]	[-0.25]	[-2.56]	[-3.23]	[0.88]	[0.39]	[0.75]	[0.86]
Panel B: Stocks with low short-sale constraints								
SSC Measures	High RIO	Low RIO	High - Low Ex. Ret	High - Low Alpha	High RIO	Low RIO	High - Low Ex. Ret	High - Low Alpha
(1) *Short Interest*	-0.16	0.34	-0.50***	-0.65***	0.16	-0.05	0.20	0.33**
	[-0.70]	[1.20]	[-3.16]	[-4.06]	[0.71]	[-0.17]	[1.23]	[2.05]
(2) *Utilization*	-0.89**	-0.85**	-0.04	-0.11	0.97**	0.61*	0.36*	0.39**
	[-2.22]	[-2.14]	[-0.22]	[-0.62]	[2.43]	[1.69]	[1.82]	[2.03]
(3) *Indicative Fee*	-0.76**	-0.77*	0.01	-0.12	0.66**	0.66*	0.01	0.07
	[-2.25]	[-1.94]	[0.04]	[-0.70]	[2.04]	[1.73]	[0.04]	[0.39]

This table shows the factor premiums and their differences across high and low RIO stocks after controlling for short-sale constraints (SSC). We use three measures of short-sale constraints: short interest, the ratio of shares held short to the total outstanding shares; utilization, the proportion of a stock's value on loan relative to its total lendable value in the borrowing market; indicative fee, the borrowing cost of equities in the borrowing market. At the end of each month t, stocks are sequentially sorted into $2 \times 2 \times 5$ groups based on one of the SSC measures, residual institutional ownership levels $\text{RIO}_{i,t-1}$ estimated from regression (2) and composite factor exposures $\beta^{comp}_{i,j,t-1}$ estimated from regression (4). For all sorting variables, we only use values up to $t-1$ to ensure data availability at the portfolio formation month t. For stocks within the same short-sale constraint level, we calculate the excess returns of high-minus-low composite beta portfolios within each RIO group as the measure of factor premiums. We also report the differences in factor premiums and their CAPM alphas between the highest and lowest RIO groups. Following Nagel (2005), equal-weighted portfolio returns are used to mitigate the size effect. The sample period spans from January 1980 to December 2021 for portfolios based on short interest, and from May 2002 to December 2021 for portfolios based on the other two measures. The reported t-statistics use Newey-West standard errors, adjusting for serial correlations.

Table 12: **Factor Premium Differences with Control of Noise Trading**

	(A) *Characteristic*				(B) *Macro-related*			
Panel A: Controlling for noise trading risk exposures								
	High	Low	High - Low		High	Low	High - Low	
	RIO	RIO	Ex. Ret	Alpha	RIO	RIO	Ex. Ret	Alpha
High β^{FIT}	-1.24	-1.07	-0.17	-0.88*	1.3	0.69	0.61	0.90*
	[-1.24]	[-0.84]	[-0.33]	[-1.66]	[1.20]	[0.65]	[1.05]	[1.77]
Low β^{FIT}	-1.48	-0.71	-0.77	-1.30***	1.72*	0.74	0.98*	1.28**
	[-1.52]	[-0.69]	[-1.53]	[-2.75]	[1.93]	[0.80]	[1.76]	[2.20]
Panel B: Controlling for noise trading shocks								
	High	Low	High - Low		High	Low	High - Low	
	RIO	RIO	Ex. Ret	Alpha	RIO	RIO	Ex. Ret	Alpha
High FIT	-0.85	-0.41	-0.44	-1.17***	1.2	0.4	0.8	1.15**
	[-0.95]	[-0.37]	[-0.89]	[-2.58]	[1.35]	[0.42]	[1.54]	[2.56]
Low FIT	-1.05	0.41	-1.46***	-2.06***	1.47	-0.22	1.69***	1.93***
	[-1.10]	[0.38]	[-3.03]	[-4.78]	[1.60]	[-0.23]	[3.04]	[3.78]

This table shows the factor premiums and their differences across high and low RIO stocks after controlling for noise trading exposures (Panel A) or noise trading shocks (Panel B). We use the measure of mutual fund flow-induced-trading (FIT) as noise trading shocks $\text{FIT}_{i,t}$ following equation (9), and we calculate the exposures of individual stock FIT on market FIT as the noise trading risk exposures $\beta^{\text{FIT}}_{i,t}$ following equation (10). At the end of each month t, stocks are sequentially sorted into $2 \times 2 \times 5$ groups based on noise trading risk exposures $\beta^{\text{FIT}}_{i,t}$ or noise trading shocks $\text{FIT}_{i,t}$, residual institutional ownership levels $\text{RIO}_{i,t-1}$ estimated from regression (2) and composite factor exposures $\beta^{comp}_{i,j,t-1}$ estimated from regression (4). For all sorting variables, we only use values up to $t-1$ to ensure data availability at the portfolio formation month t. For stocks within the same β^{FIT} or FIT level, we calculate the excess returns of high-minus-low composite beta portfolios within each RIO group as the measure of factor premiums. We also report the differences in factor premiums and their CAPM alphas between the highest and lowest RIO groups. Following Nagel (2005), equal-weighted portfolio returns are used to mitigate the size effect. The sample period spans from 1980Q1 to 2021Q4. The reported t-statistics use Newey-West standard errors, adjusting for serial correlations.

Appendix

Appendix A: Factor Description

Here we provide the detailed description of each factor, as well as the data sources.

A.1 Characteristic Factors

1 & 2. Size (SMB) and Book-to-Market Ratio (HML) Small-size stocks or with a high book-to-market ratio earn higher returns—a finding documented in Fama and French (1992) and subsequently widely acknowledged. In Fama and French (1993), SMB (small minus big) and HML (high minus low) are constructed as factors capturing size and book-to-market risks that are not incorporated by the aggregate market return. Monthly SMB and HML factors are downloaded from French's library.

3. Momentum (MOM) The momentum factor is constructed by longing stocks with good past performance and shorting stocks with bad performance, a widely known strategy applied in industry. This performance "persistence" is documented in Jegadeesh and Titman (1993), showing that stocks with high past 2- to 12-month returns will generate higher returns in the future. The monthly Mom factor is downloaded from French's library.

4 & 5. Investment (CMA) and Profitability (RMW) In Fama and French (2015), factors based on investment and operating profitability are proposed, and the original Fama-French three factors are expanded into five factors. Firms with high investments (aggressive) earn lower returns than firms with low investments (conservative), and factor CMA (conservative minus aggressive) are constructed by longing low investment stocks and short-selling high investment stocks. For profitability, firms with robust profitability deliver higher returns than firms with weak profitability, and factor RMW (robust minus weak) are constructed accordingly. Monthly CMA and RMW are also downloaded from French's library.

6 & 7 & 8. Investment-on-Assets (R_{IA}), ROE (R_{ROE}) and Expected Growth (R_{EG}) These three factors are from the investment-CAPM theory proposed by Hou et al. (2015) and Hou et al. (2021). They find that stocks with lower investments, with higher profitability measured by ROE, or higher 1-year expected investment growth will deliver higher expected returns. The corresponding factors R_{IA}, R_{ROE}, R_{EG}, along with market return (R_{MKT}) and size premium factors (R_{ME}) are q^5 model and can explain large cross-sectional variations in individual stocks' expected returns. Monthly R_{IA}, R_{ROE} , R_{EG} are downloaded from Zhang Lu's Website [21].

9. Gross Profitability (PMU) Novy-Marx (2013) documents that sorting on the ratio of gross profitability to total assets delivers higher returns for more profitable firms, which is

[21] http://global-q.org/factors.html

difficult to reconcile with value premium explanations and can digest most earnings-related (or some of the unrelated) anomalies. This paper reconstructs and expands PMU factors up until 2021 following Novy-Marx (2013) (96% correlation with original PMU factors). All stocks are independently divided into 2×3 value-weighted size/gross profitability (GP) portfolios using NYSE breakpoints(median for size, 30th/70th for GP) at the end of each June, where GP is defined as the total revenues minus cost of goods sold divided by total assets. The PMU factor is obtained by longing small and big most profitable portfolios while shorting small and big most unprofitable portfolios.

10 & 11. Managements (MGMT) and Performance (PERF) Stambaugh and Yuan (2017) propose two mispricing factors based on the 11 well-documented anomalies examined by Stambaugh et al. (2012), Stambaugh et al. (2014), and Stambaugh et al. (2015). The MGMT factor comes from the first class of 6 anomalies that are directly affected by firm's management. The PERF factor comes from the second class of 5 anomalies that are more related to performance and less likely to be affected by managers. These two mispricing factors are reconstructed and expanded to 2021 following the appendix of Stambaugh and Yuan (2017) (98% and 95% correlation to original MGMT and PERF factors). Factors are long-short portfolios based on the size and average mispricing score of the two clusters of mispricing measures.

12. Quality (QMJ) Asness et al. (2019) define "quality" as the firm's composite characteristic for which investors are willing to pay a higher price. The measure of quality contains 21 firm-level characteristics along three dimensions: profitability, growth and safety. The quality-minus-junk (QMJ) factor that longs high-quality stocks and shorts low-quality stocks earns significant risk-adjusted returns in the United States and across 24 countries, as documented in Asness et al. (2019), and is one of the most important factors used in and provided by AQR. The QMJ factor is downloaded from the AQR website.[22]

A.2 Macro-Related Factors

1. Consumption (CON) Consumption-based asset pricing is the classic asset pricing model that was first introduced by Lucas (1978). Often referred to as the "Lucas tree," it indicates that an asset should have a higher expected return only if it covaries more with consumption growth. However, numerous studies have proved the failure of CCAPM. See, for example, Mehra and Prescott (1985), Chen et al. (1986). The CON factor is the growth of the real personal consumption expenditure on non-durable goods and services per capita. Monthly personal consumption data are obtained from the Bureau of Economic Analysis (BEA) Table 2.8.5. Monthly implicit price deflators data (for real consumption adjustment) are from BEA Table 1.1.9. Monthly population data are from BEA Table 2.6.

[22] https://www.aqr.com/Insights/Datasets/Quality-Minus-Junk-Factors-Monthly

2. Total Factor Productivity (TFP) In standard production-based asset pricing models(e.g. Jermann (1998)), TFP growth should be positively priced. Firms with a large degree of exposure to TFP growth should have higher expected returns since they will perform badly during recessions. Quarterly TFP data are from the website of the Federal Reserve Bank of San Francisco.

3. Industrial Production (IPG) Similar to TFP, industrial production growth is also a proxy for an aggregate productivity shock, which should also be positively priced. Monthly IPG data are from the Federal Reserve Bank of St. Louis FRED website.

4. Default Spread (DEF) The default spread is defined as the difference between BAA bond yields and AAA bond yields, which is widely believed to be countercyclical (see, e.g., Gilchrist and Zakrajšek (2012)). To avoid forward looking, We use first-order differences instead of AR(1) or other econometric specifications of the default spread as an innovation measure; thus, the corresponding DEF factor is the change in the default spread. Monthly BAA and AAA data are from the Federal Reserve Bank of St. Louis FRED website.

5. Term Spread (TERM) The term spread (yield curve) is calculated as the difference between the 10-year Treasury bond yield and the 1-year Treasury bond yield. An inverted yield curve is often regarded as a signal of a recession, and thus the term spread is more likely to be pro-cyclical. For example, Campbell Harvey's dissertation (Harvey (1986)) showed that an inverted yield curve accurately forecasts U.S. recessions. The TERM factor is also calculated as the change in the term spread. Monthly 10-year and 1-year Treasury data are from the Federal Reserve Bank of St. Louis FRED website.

6 & 7. Unexpected Inflation (UI) and Change in Expected Inflation (DEI) Inflation is also a pervasive macro variable and is closely related to economic conditions. Following Chen et al. (1986) and Shen et al. (2017), inflation can be divided into unanticipated inflation $UI_t = I_t - E_{t-1}[I_t]$ and change in expected inflation $DEI_t = E_t[I_{t+1}] - E_{t-1}[I_t]$, where $I_t \equiv (CPI_t - CPI_{t-1})/CPI_{t-1}$. The expectation estimation method follows Fama and Gibbons (1984), where the change in inflation $I_t - I_{t-1}$ is modelled as an MA(1) process. The monthly CPI data are from Federal Reserve Bank of St. Louis FRED website.

8. Market Excess Return (MKT) CAPM is regarded as one of the most important classic asset pricing models. Strictly speaking, the market return is not a non-equity macro variable, but it has a very strong economic foundation and has gained great prominence in both the finance academy and industry. Thus, the market return is included as one of the macro-related factors that are believed to contain true risks. Many possible forces that cause an empirical failure of CAPM have been proposed by previous work. See, for example, leverage constraint (Frazzini and Pedersen (2014)), disagreement (Hong and Sraer (2016)), and investor sentiment(Shen et al. (2017)).

9. Aggregate Market Volatility (VOL) Market usually goes to expansion when volatility drops; thus, the change in aggregate volatility should be negatively priced as a

countercyclical factor. Stocks hedging for market volatility will be attractive to risk-averse investors. The high demand for stocks with a high degree of exposure to a change in volatility will push down the price and push up the stocks' expected return. Chen et al. (2011) have explored in detail volatility's prediction of market return. Here, We examine the risk premium of realized monthly volatility using the SP500 index daily return following French et al. (1987), the change of which is used as volatility factor.

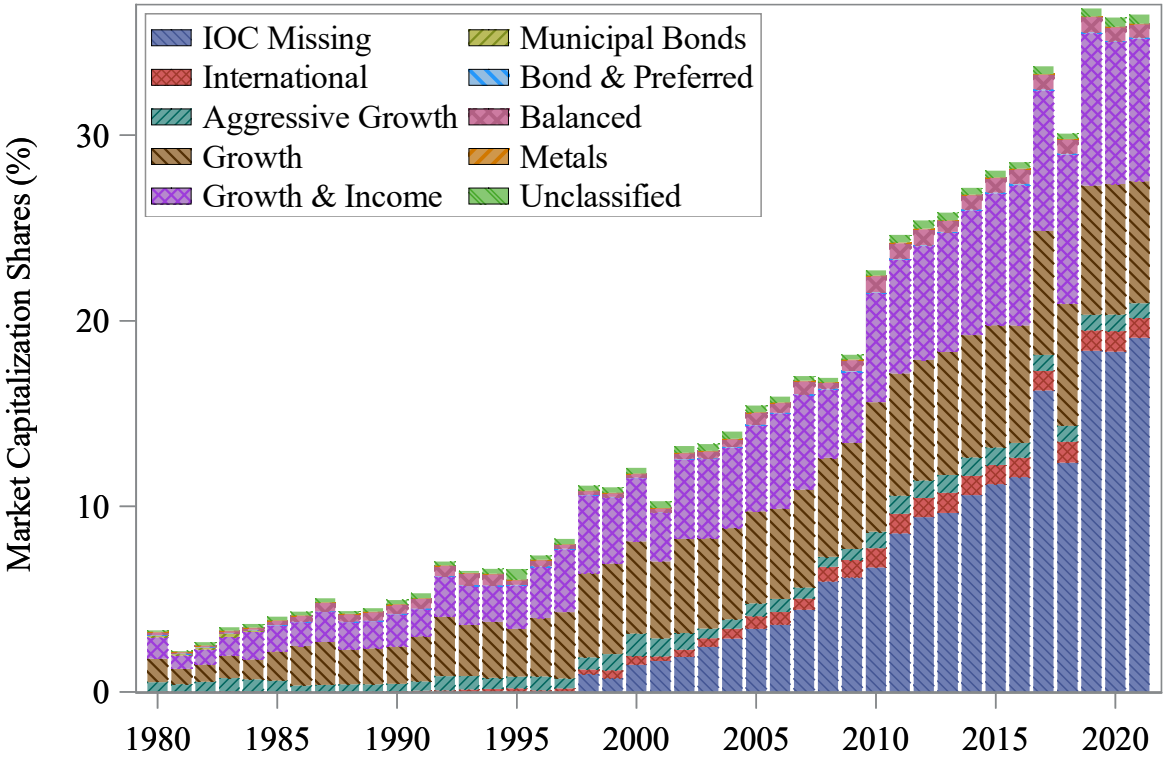
10. Labor Income (LAB) As a measure of human capital, labor income has been introduced in the conditional CAPM model (e.g.,Jagannathan and Wang (1996)) and has also proved to be positively priced cross-sectionally (e.g.,Santos and Veronesi (2005). Labor income is calculated as the difference between total personal income and dividend income per capita following Jagannathan and Wang (1996) , and the LAB factor is labor income growth. Labor income and population data are collected from the Bureau of Economic Analysis (BEA) National Table 2.6.

Appendix B: Mutual Fund Holdings

We present a comprehensive categorization of mutual fund holding sbased on the Investment Objective Code (IOC) supplied by the Thomson Reuters Mutual Fund Holdings (S12) Database. Figure B1 illustrates the market capital percentages for nine distinct mutual fund classes, as well as those without IOC records (IOC Missing). The figure reveals a significant increase in the total market capitalization of mutual funds from 1980 to 2021. Growth and growth & income investment objective funds dominate nearly the entire mutual fund market before 2000 and account for approximately half of the market post-2000. In the period after 2000, funds with missing IOC data experience rapid asset growth. By the end of 2021, the market capitalization percentages for IOC missing funds, growth funds, and growth & income funds stand at 19.1%, 6.6%, and 7.7%, respectively.

The remarkable expansion of IOC missing funds mirrors the rise of indexing funds. Additionally, Table B1 displays the names and assets of the top 10 mutual funds with missing IOC at the end of 2021. The largest IOC missing fund is the well-known SPY ETF, succeeded by various index or passive funds provided by prominent financial institutions such as Vanguard and BlackRock. Consequently, it is reasonable to infer that IOC missing funds are predominantly indexing funds.

Figure B1: **Mutual Fund Holdings Decomposition**



This figure shows the annual market capitalization shares of mutual funds types based on investment objective codes (IOC) provided by Thomson Reuters Mutual Fund Holdings (S12) Database from WRDS.

Table B1: **Top 10 Largest MF with missing IOC**

	Assets(US $1Mil)	FundName(Abbrev.)
1	416,901	SPDR S&P 500 ETF
2	389,783	VANGUARD TOT INTL STK IN
3	332,656	ISHARES S&P 500 INDEX FU
4	299,607	VANGUARD TOT BD MKT INDE
5	245,861	VANGUARD TOTAL BOND MARK
6	198,802	POWERSHARES NASDAQ 100 S
7	163,488	VANGUARD TAX-MANAGED INT
8	130,818	PIMCO INCOME FUND
9	127,389	CREF STOCK ACCOUNT
10	100,257	ISHARES CORE MSCI EAFE E

This table shows the largest top 10 US mutual funds with missing IOC records in 2021.12 from WRDS Thomson Reuters Mutual Fund Holdings (S12) Database.

Appendix C: Supplementary Results

C.1 Results of Value-weighted Portfolios

To mitigate size effects on the beta-return relations, we use equal-weighted portfolios in the paper. This method may potentially put biased high weights to small stocks. In this section, we present our main results using value-weighted methods, showing that our findings are robust regardless of weighting schemes.

Specifically, we reproduce the main results in Table 4 under the log-value-weighted and value-weighted portfolios. Contrary to equal-weihghted methods, some studies argues that the value-weighted methods may put too much weights on mega firms, eg. Belo et al. (2017) use employee numbers weighted portfolios and Jensen et al. (2023) use capped value weighted portfolios. Thus, besides the standard value-weighted methods, we also use log-value-weighted methods in our analysis. At the end of every month t, stocks are independently double-sorted into 5×5 groups based on their residual institutional ownership levels $\text{RIO}_{i,t-1}$ estimated from regression (2) and composite factor exposures $\beta^{comp}_{i,j,t-1}$ estimated from regression (4). For both sorting variables, we only use values up to $t-1$ to ensure data availability at the portfolio formation month t. Table C1 shows the main results under log-value-weighted and value-weighted methods, which is in the same format as Table 4 showing the excess returns and CAPM α for double-sorted portfolios with highest and lowest β and various levels of RIO.

Our results remain consistent when using log-value-weighted and value-weighted portfolios. The factor premiums, measured by the excess returns of high-minus-low beta-sorted portfolios, are -0.61% lower (t-stat=-2.72) and -0.54% lower (t-stat=-2.04) for characteristic factors, while they are 0.48% higher (t-stat=2.34) and 0.37% higher (t-stat=1.49) for macro-related factors in high RIO stocks. These findings become even stronger when we control for market exposures. The CAPM α differences in factor premiums between high and low RIO stocks are -0.80% lower and -0.69% lower for characteristic factors, and 0.61% higher and 0.53% higher for macro-related factors, all statistically significant at the 5% confidence level.

It is worth noting that the differences in stock returns due to RIO levels are less significant under value-weighted methods. This suggests that small stocks are more sensitive to demand shocks from institutions, as described by Gompers and Metrick (2001), thus the higher returns from high RIO are mostly driven by small stocks. Nevertheless, it didn't affect our findings of factor premium differences as our focus is on examining the effects of RIO on factor premiums rather than on stock returns.

In summary, the our results are robust and significant under value-weighted schemes, indicating that our findings are reliable and not solely driven by micro-cap stocks.

Table C1: **Factor Premiums in Stocks with High and Low Institutional Ownerships (Value-Weighted)**

	Log value weighted						Value weighted					
Panel A. Characteristic factors												
	Low RIO	2	3	4	High RIO	High - Low	Low RIO	2	3	4	High RIO	High - Low
High β	0.62	0.88	0.94	0.93	0.76	0.14	0.82	0.81	0.91	0.85	0.61	-0.21
	[3.02]	[4.32]	[4.29]	[3.70]	[2.87]	[1.10]	[4.56]	[4.18]	[4.67]	[3.86]	[2.51]	[-1.20]
Low β	0.18	0.61	0.77	0.98	0.93	0.76***	0.39	0.63	0.69	0.59	0.72	0.33
	[0.40]	[1.48]	[1.99]	[2.49]	[2.37]	[4.12]	[0.99]	[1.62]	[1.95]	[1.70]	[1.89]	[1.49]
High - Low	0.44	0.27	0.17	-0.05	-0.18	-0.61***	0.43	0.19	0.22	0.26	-0.11	-0.54**
	[1.15]	[0.81]	[0.61]	[-0.19]	[-0.69]	[-2.72]	[1.18]	[0.57]	[0.73]	[0.97]	[-0.34]	[-2.04]
CAPM α	1.02***	0.79**	0.61**	0.38*	0.22	-0.80***	1.03***	0.74**	0.71**	0.71***	0.34	-0.69***
	[2.94]	[2.53]	[2.32]	[1.65]	[0.90]	[-3.71]	[3.15]	[2.28]	[2.37]	[2.69]	[1.07]	[-2.60]
Panel B. Macro-related factors												
	Low RIO	2	3	4	High RIO	High - Low	Low RIO	2	3	4	High RIO	High - Low
High β	0.27	0.66	0.84	1.01	1.05	0.78***	0.30	0.81	0.53	0.87	0.83	0.53**
	[0.62]	[1.70]	[2.25]	[2.58]	[2.61]	[4.62]	[0.83]	[2.39]	[1.54]	[2.51]	[2.24]	[2.28]
Low β	0.57	0.80	0.86	0.94	0.86	0.30**	0.78	0.71	0.72	0.98	0.93	0.16
	[2.83]	[3.93]	[3.96]	[3.90]	[3.11]	[2.13]	[4.26]	[4.15]	[3.79]	[4.47]	[3.66]	[0.80]
High - Low	-0.29	-0.14	-0.02	0.07	0.18	0.48**	-0.47	0.10	-0.19	-0.11	-0.10	0.37
	[-0.84]	[-0.47]	[-0.08]	[0.29]	[0.78]	[2.34]	[-1.43]	[0.35]	[-0.65]	[-0.39]	[-0.38]	[1.49]
CAPM α	-0.82***	-0.62**	-0.39*	-0.30	-0.21	0.61***	-1.00***	-0.38	-0.61**	-0.49*	-0.46*	0.53**
	[-2.59]	[-2.25]	[-1.66]	[-1.33]	[-0.98]	[2.98]	[-3.18]	[-1.42]	[-2.02]	[-1.84]	[-1.72]	[1.98]

This table shows the excess returns and CAPM α of log value weighted and value weighted double-sorted portfolios based on composite beta and residual institutional ownership. At the end of every month t, stocks are independently double-sorted into 5×5 groups based on their residual institutional ownership levels $\text{RIO}_{i,t-1}$ estimated from regression (2) and composite factor exposures $\beta^{comp}_{i,j,t-1}$ estimated from regression (4). Panel A reports the results of portfolios sorted by characteristic factor composite betas, while Panel B focuses on portfolios sorted by macro-related factor composite betas. For both sorting variables, we only use values up to $t-1$ to ensure data availability at the portfolio formation month t. Portfolio returns are log-value-weighted or value-weighted. The table shows excess returns of 10 portfolios with highest and lowest β^{comp} for each of the five RIO quintitles. High-Low rows represent excess returns for portfolios constructed by taking a long position in the highest β^{comp} portfolios and a short position in the lowest β^{comp} portfolios within each RIO quintitles. High-Low columns show return and CAPM α differences between highest and lowest RIO portfolios. The CAPM α is estimated using the regression: $r_{p,t} = \alpha_p + \beta_p \times r_{mkt,t} + \epsilon_{p,t}$ for portfolio high-minus-low p. The sample period spans from January 1980 to December 2021. The reported t-statistics use Newey-West standard errors, adjusting for serial correlations.

C.2 Results Controlling Various Factor Exposures

In this section, we report the differences in factor premiums between high and low RIO stocks after controlling for various factor model exposures. In the main results, shown in Table 4, we only report the αs adjusted by the CAPM model, which is the most theoretically grounded and widely accepted factor model. To confirm the robustness of our findings, we also show the αs after adjustments using several other factor models in this section, including the Fama-French 3-factor model (which accounts for market, size, and book-to-market ratio factors), the Carhart 4-factor model (which adds a momentum factor to the Fama-French 3-factor model), and the Fama-French 5-factor model (which includes the three factors from the Fama-French 3-factor model plus investment and profitability factors). These factor models are widely used and are also examined as characteristic factors in this study. Details about these factors can be found in Table 2 and Appendix A.1.

Table C2 presents the results after adjusting for exposures to various factor models, considering both equal-weighted and (log) value-weighted portfolios. The table reports factor premiums, measured by the excess returns of high-minus-low long-short beta-sorted portfolios, within high and low RIO stocks and their differences. The αs are the intercepts after regressing these excess returns on various factor models. The first two rows, labeled "Excess Returns" and "CAPM α", correspond to our main results shown in Table 4 and the value-weighted results in Table C1. The subsequent three rows show the αs after adjusting for the three additional factor models.

It is obvious that the differences in factor premiums remain quite robust even after adjusting for various factor models. For characteristic factors, the αs of the premium differences between high and low RIO stocks range from -0.50% to -0.85% per month, all significant at a confidence level of 10% or better across all factor models and weighting schemes. For macro-related factors, the αs of the premium differences range from 0.33% to 0.60% per month and are mostly significant. These results further prove that our findings are robust and that the observed results persist even after accounting for different factor exposures.

C.3 Results under Two-Factor Model Exposures

There has been debates about the potential bias in factor beta estimations. The one-factor model we used may not necessarily represent the "true" underlying factor structures, leading to the bias estimates. However, determining the "true" factors is an elusive task, thus choosing one-factor model is a natural approach to form portfolios with largest spreads in factor exposures. Moreover, the primary focus of this paper is to examine whether these factors have pricing powers, rather than debating whether they're the "true" factors or not.

Nonetheless, to address the concerns regarding factor beta estimation, we use a two-factor model, which incorporates the commonly recognized market factor and the factor of interest

Table C2: **Factor Premium Differences Controlling for Various Factor Models**

	Equal weighted			Log value weighted			Value weighted		
Panel A. Characteristic factors									
	Low RIO	High RIO	High - Low	Low RIO	High RIO	High - Low	Low RIO	High RIO	High - Low
Excess Return	0.32	-0.41	-0.72***	0.44	-0.18	-0.61***	0.43	-0.11	-0.54**
	[0.84]	[-1.52]	[-3.23]	[1.15]	[-0.69]	[-2.72]	[1.18]	[-0.34]	[-2.04]
CAPM α	0.86	-0.04	-0.90***	1.02	0.22	-0.80***	1.03	0.34	-0.69***
	[2.52]	[-0.15]	[-4.26]	[2.94]	[0.90]	[-3.71]	[3.15]	[1.07]	[-2.60]
FF 3-factor α	0.76	-0.09	-0.85***	0.92	0.16	-0.76***	0.96	0.23	-0.73***
	[2.59]	[-0.37]	[-4.33]	[3.19]	[0.73]	[-4.01]	[3.23]	[0.76]	[-2.64]
Carhart 4-factor α	0.32	-0.50	-0.82***	0.49*	-0.19	-0.67***	0.66	0.00	-0.66**
	[1.02]	[-1.98]	[-3.93]	[1.68]	[-0.84]	[-3.32]	[2.21]	[0.01]	[-2.50]
FF 5-factor α	0.03	-0.61	-0.64***	0.14	-0.36	-0.50***	0.13	-0.43	-0.55*
	[0.09]	[-2.69]	[-3.25]	[0.57]	[-1.94]	[-2.62]	[0.51]	[-1.72]	[-1.83]
Panel B. Macro-related factors									
	Low RIO	High RIO	High - Low	Low RIO	High RIO	High - Low	Low RIO	High RIO	High - Low
Excess Return	-0.12	0.41	0.54***	-0.29	0.18	0.48**	-0.47	-0.10	0.37
	[-0.36]	[1.63]	[2.63]	[-0.84]	[0.78]	[2.34]	[-1.43]	[-0.38]	[1.49]
CAPM α	-0.61	0.02	0.63***	-0.82	-0.21	0.61***	-1.00	-0.46	0.53**
	[-1.96]	[0.07]	[2.96]	[-2.59]	[-0.98]	[2.98]	[-3.18]	[-1.72]	[1.98]
FF 3-factor α	-0.53	0.01	0.54***	-0.74	-0.22	0.52***	-0.98	-0.41	0.56**
	[-1.91]	[0.03]	[2.62]	[-2.70]	[-1.05]	[2.89]	[-3.29]	[-1.62]	[2.13]
Carhart 4-factor α	-0.13	0.42	0.55**	-0.35	0.17	0.51***	-0.71	-0.10	0.60**
	[-0.42]	[1.81]	[2.49]	[-1.24]	[0.79]	[2.70]	[-2.23]	[-0.43]	[2.27]
FF 5-factor α	0.08	0.42	0.33	-0.10	0.20	0.30*	-0.30	0.13	0.43
	[0.28]	[1.64]	[1.64]	[-0.37]	[0.94]	[1.72]	[-1.06]	[0.59]	[1.60]

This table shows the excess returns and various factor model αs of factor premiums in high and low RIO stocks, and their differences.At the end of every month t, stocks are independently double-sorted into 5×5 groups based on their residual institutional ownership levels $\text{RIO}_{i,t-1}$ estimated from regression (2) and composite factor exposures $\beta^{comp}_{i,j,t-1}$ estimated from regression (4). For both sorting variables, we only use values up to $t-1$ to ensure data availability at the portfolio formation month t. The table shows excess returns and αs of high-minus-low portfolios constructed by longing highest β portfolios and shorting lowest β portfolios within highest and lowest RIO groups, and their differences. The CAPM α are estimated following: $r_{p,t} - r_{f,t} = \alpha_p + \beta_{p,mkt} \times (r_{mkt,t} - r_{f,t}) + \epsilon_{p,t}$ for portfolio p. The FF 3-factor α are estimated following: $r_{p,t} - r_{f,t} = \alpha_p + \beta_{p,mkt} \times (r_{mkt,t} - r_{f,t}) + \beta_{p,smb} \times r_{smb,t} + \beta_{p,hml} \times r_{hml,t} + \epsilon_{p,t}$. The Carhart 4-factor α are estimated following: $r_{p,t} - r_{f,t} = \alpha_p + \beta_{p,mkt} \times (r_{mkt,t} - r_{f,t}) + \beta_{p,smb} \times r_{smb,t} + \beta_{p,hml} \times r_{hml,t} + \beta_{p,mom} \times r_{mom,t} + \epsilon_{p,t}$. The FF 5-factor α are estimated following: $r_{p,t} - r_{f,t} = \alpha_p + \beta_{p,mkt} \times (r_{mkt,t} - r_{f,t}) + \beta_{p,smb} \times r_{smb,t} + \beta_{p,hml} \times r_{hml,t} + \beta_{p,cma} \times r_{cma,t} + \beta_{p,rmw} \times r_{rmw,t} + \epsilon_{p,t}$. The sample period is from January 1980 to December 2021. The reported t-statistics use Newey-West standard errors, adjusting for serial correlations.

k, to re-estimate the beta exposure on factor k. The beta in two-factor model, denoted as $\hat{\beta}^{pre}_{i,k,t}$ for individual stock i with respect to factor k at month t is estimated as followed.

$$r_{i,\tau} - r_{f,\tau} = \alpha_{i,k,t} + \hat{\beta}^{pre}_{i,k,t} \times f_{k,\tau} + \hat{\beta}^{pre}_{i,mkt,t} \times (r_{mkt,\tau} - r_{f,\tau}) + \epsilon_{i,k,\tau} \tag{C1}$$

Here, τ corresponds to the months within the five-year window preceding time t, $r_{i,\tau}$ is the return of stock i at month τ, $r_{f,\tau}$ is the risk-free rate at month τ, $f_{k,\tau}$ is the value of factor k at month τ, and $r_{mkt,\tau}$ is the market return at month τ. To create composite beta $\hat{\beta}^{comp}_{i,j,t}$, we follow the process described in section 2.2.2, only replacing the one-factor pre-formation beta with the newly estimated two-factor beta $\hat{\beta}^{pre}_{i,k,t}$.

Table C3 reports the results based on two-factor composite beta. We follow the same portfolio analysis process describes in section 3.1 to form double-sorted portfolios based on RIO and composite beta, only replacing the one-factor composite beta $\beta^{comp}_{i,j,t}$ to two-factor beta $\hat{\beta}^{comp}_{i,j,t}$. Table C3 is in the same format of main results in Table 4. The high-minus-low rows show the return spreads between highest and lowest beta portfolios, estimating the factor premiums under different RIO levels. The intersections between high-minus-low rows and columns quantify the factor premium differences between highest and lowest RIO groups

The results in Table C3 are robust and consistent with our findings. Characteristic factors have higher premiums in low RIO stocks, while macro-related factor have higher premiums in high RIO stocks. Specifically, for characteristic factors, the factor premiums decrease from 0.31% per month in lowest RIO groups to -0.31% in highest RIO groups, generating a significant difference of -0.61% (t-stat=-2.77) per month. On the contrary, macro-related factor premiums are much higher in high RIO stocks (0.37%) than those in low RIO stocks (-0.18%), leading to a positive spreads of 0.56% per month (t-stat=3.00). The results are robust and even more significant when controlling for market risk exposures. The CAPM α of factor premium differences are -0.81% (t-stat=-3.85) for characteristic factors and 0.65% (t-stat=3.47) for macro-related factors. In summary, the results are robust and insensitive to changes of beta estimation methods.

C.4 Fama-Macbeth Regressions with Institutional Ownership and Firm Size

In this section, we present the results of Fama-MacBeth regressions while directly controlling for size and other firm characteristics that may influence the relationship between beta and

Table C3: **Factor Premiums under Two-Factor Model**

	Low RIO	2	3	4	High RIO	High - Low
Panel A. Characteristic factors						
High β	0.62	0.94	1.06	1.16	1.07	0.45***
	[2.64]	[4.08]	[4.13]	[3.88]	[3.34]	[2.85]
Low β	0.31	0.80	0.82	1.22	1.37	1.06***
	[0.74]	[1.98]	[2.12]	[3.09]	[3.42]	[5.88]
High - Low	0.31	0.14	0.24	-0.07	-0.31	-0.61***
	[0.94]	[0.46]	[0.96]	[-0.31]	[-1.41]	[-2.77]
CAPM α	0.68**	0.48	0.51**	0.16	-0.14	-0.81***
	[2.13]	[1.56]	[2.02]	[0.70]	[-0.60]	[-3.85]
Panel B. Macro-related factors						
	Low RIO	2	3	4	High RIO	High - Low
High β	0.44	0.81	0.99	1.21	1.49	1.06***
	[1.10]	[2.19]	[2.74]	[3.09]	[3.66]	[5.68]
Low β	0.62	0.90	1.00	1.15	1.12	0.50***
	[2.48]	[3.51]	[3.77]	[3.73]	[3.48]	[3.47]
High - Low	-0.18	-0.08	-0.01	0.06	0.37**	0.56***
	[-0.75]	[-0.37]	[-0.05]	[0.33]	[2.01]	[3.00]
CAPM α	-0.44*	-0.35	-0.18	-0.05	0.21	0.65***
	[-1.77]	[-1.51]	[-0.91]	[-0.26]	[1.19]	[3.47]

This table shows the excess returns and CAPM α of double-sorted portfolios based on two-factor composite beta and residual institutional ownership. At the end of every month t, stocks are independently double-sorted into 5×5 groups based on their residual institutional ownership levels $\text{RIO}_{i,t-1}$ estimated from regression (2) and two-factor composite factor exposure $\hat{\beta}^{comp}_{i,j,t-1}$, the average cross-sectional percentiles of single factor betas $\hat{\beta}^{pre}_{i,k,t-1}$ as estimated in regression Equation (C1). Panel A reports the results of portfolios sorted by characteristic factor composite betas, while Panel B focuses on portfolios sorted by macro-related factor composite betas. For both sorting variables, we only use values up to $t-1$ to ensure data availability at the portfolio formation month t. Following Nagel (2005), we report equal-weighted portfolio returns to mitigate the size effect. The table shows excess returns of 10 portfolios with highest and lowest β^{comp} for each of the five RIO quintitles. High-Low rows represent excess returns for portfolios constructed by taking a long position in the highest β^{comp} portfolios and a short position in the lowest β^{comp} portfolios within each RIO quintitles. High-Low columns show return and CAPM α differences between highest and lowest RIO portfolios. The CAPM α is estimated using the regression: $r_{p,t}-r_{f,t} = \alpha_p+\beta_p\times(r_{mkt,t}-r_{f,t})+\epsilon_{p,t}$ for high-minus-low portfolio p. The sample period spans from January 1980 to December 2021. The reported t-statistics use Newey-West standard errors, adjusting for serial correlations.

returns. The Fama-MacBeth regressions are conducted as follows.

$$\begin{aligned} r_{i,t+1} - r_{f,t+1} \;=\; & a_t + b_{1,t}\ \beta^{post}_{i,k,t} \times \mathrm{IO}_{i,t} + b_{2,t}\ \beta^{post}_{i,k,t} \times ln(\mathrm{ME}_{i,t}) + b_{3,t}\ \beta^{post}_{i,k,t} \times ln^2(\mathrm{ME}_{i,t}) \\ & + b_t\ \beta^{post}_{i,k,t} \times \mathrm{control}_{i,t} + c_{1,t}\ \beta^{post}_{i,k,t} + c_{2,t}\ \mathrm{IO}_{i,t} + c_{3,t}\ ln(\mathrm{ME}_{i,t}) \\ & + c_{4,t}\ ln^2(\mathrm{ME}_{i,t}) + c_t\ \mathrm{control}_{i,t} + \epsilon_{i,t+1} \end{aligned} \tag{C2}$$

Here, $\mathrm{IO}_{i,t}$ is the institutional ownership level of stock i in month t. The most recent quarterly IO data is used. Instead of using RIO, we use IO here and mitigate the size effects by directly controlling for log of market equity $ln(\mathrm{ME}_{i,t})$ and its square $ln^2(\mathrm{ME}_{i,t})$, as used for calculating RIO in equation (2), and their interactions with beta in the regressions. Except for size, we also include other firm characteristics that may also affect the beta-return dynamics and relate to IO as control variables, including book-to-market ratio, turnover, volatility and past returns. [23] To ensure a fair comparison across different factors, both firm characteristics and post-formation betas are standardized to have a cross-sectional mean of 0 and a standard deviation of 1 prior to forming the interaction terms. These interaction terms are subsequently calculated as the product of the standardized firm characteristics and standardized betas.

Table C4 shows a robust negative (positive) partial effects of IO on characteristic (macro-related) factor beta-return dynamics. Compared to Fama-Macbeth regressions in Table 6, the impacts from IO are more significant after purging out the effects from size and other firm characteristics. When there is a 1% increase of IO, the beta-return relations weaken for 11 out of 12 characteristic factors, as the monthly excess returns response 0.53% less to one standard deviation beta increase on average. On the contrary, 8 out of 10 macro-related factors' beta-return relations are strengthened by institutional holdings. The increase of 1% IO is associated with 0.47% higher monthly excess returns as the macro-related factor beta on average increases by 1 standard deviation.

The effects of size on beta-return relations are nonlinear as $\beta \times \mathrm{ME}$ (b_2) and $\beta \times \mathrm{ME}^2$ (b_3) are significant in opposite directions. The results indicate that characteristic factors have lower premiums in small stocks, while macro-related factors have higher premiums. The results show that although size and institutional ownership are highly correlated, their impacts on factor pricings manifest in opposing directions. One potential explanation is that the size, after adjusted for IO, may serve as a proxy for disagreement levels or retail investor attentions which lead to higher limits to arbitrage and thus larger mispricings, especially the overpricing due to short-sale constraints. This potential mechanism follows the concepts of

[23] The book-to-market ratio is calculated by dividing the most recent book equity by the month-end market value. The turnover rate is estimated by dividing the monthly trading volume by the total shares outstanding. The monthly volatility is measured as the daily return volatility for each month. The past return of a stock represents the past 12- to 3-month return, estimated as the cumulative return from the previous year to the most recent quarter (t-12 to t-3 return).

Table C4: **Fama-Macbeth Regressions with Institutional Ownership and Firm Size**

		(A) Characteristic factors				*(B) Macro-related factors*		
		$\beta \times IO$	$\beta \times ME$	$\beta \times ME^2$		$\beta \times IO$	$\beta \times ME$	$\beta \times ME^2$
	COMP	-0.53***	1.00	-0.83	COMP	0.47***	-0.77	0.63
		[-3.99]	[7.36]	[-6.97]		[4.51]	[-5.39]	[5.30]
1	SMB	0.28**	-0.80	0.78	CON	0.14*	-0.15	0.09
		[2.27]	[-5.79]	[5.88]		[1.80]	[-1.39]	[1.04]
2	HML	-0.27*	0.41	-0.29	TFP	0.21***	-0.49	0.43
		[-1.88]	[2.88]	[-2.47]		[2.78]	[-4.44]	[4.56]
3	MOM	-0.27**	0.82	-0.82	IPG	0.42***	-0.70	0.65
		[-2.35]	[6.17]	[-6.46]		[4.55]	[-5.52]	[5.93]
4	CMA	-0.26*	0.50	-0.41	DEF	0.34***	-0.83	0.80
		[-1.78]	[3.78]	[-3.72]		[3.03]	[-5.36]	[5.59]
5	RMW	-0.38***	0.86	-0.78	TERM	0.24**	-0.51	0.49
		[-2.83]	[6.13]	[-6.01]		[2.41]	[-4.32]	[4.30]
6	R_{IA}	-0.22	0.56	-0.48	UI	0.33***	-0.55	0.51
		[-1.50]	[3.80]	[-3.86]		[2.93]	[-4.55]	[4.76]
7	R_{ROE}	-0.27**	0.76	-0.71	DEI	-0.29**	0.88	-0.86
		[-2.18]	[5.68]	[-5.85]		[-2.52]	[6.74]	[-6.80]
8	R_{EG}	-0.34***	0.78	-0.71	MKT	0.31***	-0.54	0.38
		[-2.60]	[6.36]	[-6.36]		[2.79]	[-3.94]	[3.11]
9	PMU	-0.32***	0.52	-0.45	VOL	0.35***	-0.73	0.66
		[-2.76]	[4.14]	[-4.01]		[2.85]	[-5.15]	[5.08]
10	MGMT	-0.31**	0.82	-0.70	LAB	-0.15	0.56	-0.60
		[-2.07]	[6.02]	[-5.94]		[-1.28]	[4.70]	[-5.22]
11	PERF	-0.33***	0.73	-0.66				
		[-2.90]	[5.40]	[-5.29]				
12	QMJ	-0.47***	1.02	-0.91				
		[-3.62]	[8.11]	[-7.76]				

This table shows the results of Fama-MacBeth regressions based on equation (C2), reporting the time-series averages of the interaction coefficients b_1 ($\beta \times IO$), b_2 ($\beta \times ME$) and b_3 ($\beta \times ME^2$) for each factor. The post-formation beta, $\beta^{post}_{i,k,t}$, for stock i on factor k in month t, is defined as the post-formation beta, $\beta^{post}_{p,k}$, of the pre-formation beta-sorted portfolio p (as estimated using equation (5)) to which stock i belongs in month t. The composite post-formation betas are calculated as the average percentile rankings of single-factor post-formation betas. Control variables are firm characteristics including log market capitalization, square of log market capitalization, book-to-market ratio, turnover, volatility and past returns. Both firm characteristics and post-formation betas are standardized to have a cross-sectional mean of 0 and a standard deviation of 1 prior to forming the interaction terms. The sample period spans from January 1980 to December 2021. The reported t-statistics use Newey-West standard errors, adjusting for serial correlations.

Hong and Sraer (2016) and Da et al. (2011) in the cross-sectional dimension. The observed significant impacts of size on beta-return relations underscore the necessity of controlling for size in our analysis.

C.5 Combined Effects of RIO and Other Time-Series Variables

In section 4.3, we examine the combined effects of RIO and sentiment on factor premiums to highlight the interaction between cross-sectional and time-series forces. In this section, we demonstrate the robustness and consistency of our results by showing that the combined effects also hold when analyzing other time-series variables. Specifically, we investigate three additional time-series forces discussed in section 4.2: inflation changes, broker-dealer leverage, and investor disagreement. Details regarding the data sources and construction of these time-series variables are provided in section 4.2.

To examine the joint impact of cross-sectional and time-series dimensions, we analyze factor premiums within stocks grouped by RIO and across different subsample periods, split by the levels of the time-series variables discussed above. We use the same methodology as in section 4.3 to construct double-sorted portfolios based on RIO and composite factor betas, with the only difference being that sentiment is replaced by one of the aforementioned time-series variables.

Table C5 reports the characteristic and macro-related factor premiums across high and low RIO groups and different subsample periods. The results show that characteristic factors are most prominent within low RIO groups and during "low rationality" periods, i.e., periods of high inflation changes, high broker-dealer leverage, or high disagreement, where characteristic factor premiums reach 0.77%, 1.20%, and 0.27%, respectively. In contrast, macro-related factors show the largest premiums within high RIO stocks and during "high rationality" periods, i.e., periods of low inflation changes (0.86%), low leverage ratios (0.87%), and low disagreement (0.44%).[24] These results complement the analysis of sentiment and RIO, further supporting our findings on the combined explanatory power of time-series and cross-sectional dimensions on factor premiums.

[24]Notably, the first two time-series variables have a stronger impact on factor premiums than investor disagreement, consistent with the regression results discussed in Table 8.

Table C5: **Combined Effects of Time-Series Variables and RIO**

	(A) Characteristic factors				*(B) Macro-related factors*		
Panel A. High and low ΔInflation periods							
	Low RIO	High RIO	High - Low		Low RIO	High RIO	High - Low
High ΔInf	0.77	0.03	-0.74**	High ΔInf	-0.57	-0.04	0.53*
	[1.46]	[0.07]	[-2.36]		[-1.14]	[-0.10]	[1.88]
Low ΔInf	-0.13	-0.84**	-0.71**	Low ΔInf	0.31	0.86**	0.54**
	[-0.26]	[-2.24]	[-2.34]		[0.70]	[2.42]	[2.00]
High - Low	0.9	0.86	-0.03	High - Low	-0.88	-0.90*	-0.02
	[1.25]	[1.63]	[-0.07]		[-1.30]	[-1.82]	[-0.04]
Panel B. High and low broker-dealer leverage periods							
	Low RIO	High RIO	High - Low		Low RIO	High RIO	High - Low
High Leverage	1.20**	0.39	-0.82***	High Leverage	-0.87*	-0.14	0.73***
	[2.49]	[1.04]	[-3.12]		[-1.79]	[-0.43]	[2.66]
Low Leverage	-0.41	-1.06**	-0.65**	Low Leverage	0.49	0.87**	0.38
	[-0.77]	[-2.55]	[-2.02]		[1.06]	[2.23]	[1.33]
High - Low	1.61**	1.44***	-0.17	High - Low	-1.35**	-1.00**	0.35
	[2.33]	[2.65]	[-0.42]		[-2.18]	[-2.01]	[0.94]
Panel C. High and low aggregate disagreement periods							
	Low RIO	High RIO	High - Low		Low RIO	High RIO	High - Low
High Disagree	0.27	-0.59	-0.85**	High Disagree	-0.18	0.46	0.64*
	[0.41]	[-1.32]	[-2.42]		[-0.31]	[1.09]	[1.93]
Low Disagree	0.36	-0.24	-0.60**	Low Disagree	-0.07	0.37	0.44*
	[0.98]	[-0.75]	[-2.44]		[-0.19]	[1.25]	[1.80]
High - Low	-0.09	-0.34	-0.25	High - Low	-0.12	0.09	0.21
	[-0.12]	[-0.60]	[-0.60]		[-0.17]	[0.17]	[0.51]

This table shows the factor premiums within stocks grouped by RIO levels and across different sample periods. At the end of each month t, stocks are double-sorted into 5×5 groups based on their residual institutional ownership levels $\text{RIO}_{i,t-1}$ estimated from regression (2) and composite factor exposures $\beta^{comp}_{i,j,t-1}$ estimated from regression (4). For both sorting variables, we only use values up to $t-1$ to ensure data availability at the portfolio formation month t. Factor premiums are calculated as the excess returns of high-minus-low beta-sorted portfolios within each RIO tier. Following Nagel (2005), we report equal-weighted portfolio returns to mitigate the size effect. The sample is divided into two periods based on one of the following three time-series variables, using their median levels over the full sample: (1) changes in monthly inflation, calculated using Consumer Price Index (CPI) data from the Federal Reserve Bank of St. Louis FRED database; (2) the broker-dealer leverage measure, as defined by Adrian et al. (2014) and obtained from Tyler Muir's website, with data available through the third quarter of 2017; and (3) investor disagreement, as defined by Yu (2011), estimated using the aggregate standard deviations of long-term earnings-per-share forecasts from the IBES database, with data starting in December 1981. Panels A, B, and C report the characteristic and macro-related factor premiums across different RIO tiers following the subsample periods split by one of the three time-series variables. The reported t-statistics use Newey-West standard errors, adjusting for serial correlations.