

Dispersed Ownership and Asset Pricing: An Unpriced Premium Associated with Free Float

Finance Working Paper N° 826/2022

April 2025

Bruce Hearn
University of Southampton

Igor Filatotchev
King's College London

Marc Goergen
IE Business School and ECGI

© Bruce Hearn, Igor Filatotchev and Marc Goergen 2025. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

This paper can be downloaded without charge from:
http://ssrn.com/abstract_id=4102188

www.ecgi.global/content/working-papers

ECGI Working Paper Series in Finance

Dispersed Ownership and Asset Pricing: An Unpriced Premium Associated with Free Float

Working Paper N° 826/2022

April 2025

Bruce Hearn
Igor Filatotchev
Marc Goergen

The authors are grateful for helpful comments received from participants at the 2023 annual meetings of the International Corporate Governance Society at IE Business School, Madrid, Spain. The research reported in this paper was partially funded by the Spanish State Research Agency (AEI) /10.13039/501100011033) – Grant No. PID2020-115982RB-C2 and Grant No. PID2023148086NB-C22 granted by MCIU /AEI /10.13039/501100011033 / FEDER, UE.

© Bruce Hearn, Igor Filatotchev and Marc Goergen 2025. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Abstract

We explore differences in the levels of dispersed ownership that lead to a returns-based free float hedging factor in addition to size, which augments the capital asset pricing model (CAPM) in explaining the cross-section of stock returns. Using the S&P 1500 stocks in the US between 1985 and 2023, the results support the advantages of free float within a three-factor CAPM including size over alternative models based on liquidity, book-to-market value, and momentum. We argue that this yields a useful means for hedging effectively against the risks associated with the fundamental underlying likelihood of expropriation in a specific firm based on its ownership structure.

Keywords: CAPM, free float, ownership and control, investor protection

JEL Classifications: G11, G12, G15, O55

Bruce Hearn

Professor of Governance
University of Southampton
University Road
Southampton SO17 1BJ, UK
e-mail: B.A.Hearn@soton.ac.uk

Igor Filatotchev

Professor of Corporate Governance and Strategy
King's College London
30 Aldwych
London WC2B 4BG, UK
phone: +44 2078365454
e-mail: igor.filatotchev@kcl.ac.uk

Marc Goergen*

Professor of Finance
IE Business School, Department of Finance
María de Molina 11
28006 Madrid, Spain
phone: +34 917 875 182
e-mail: marc.goergen@ie.edu

*Corresponding Author

Dispersed ownership and asset pricing: An unpriced premium associated with free float

Bruce Hearn

Department of Accounting, Southampton Business School, University of Southampton, UK
Email: b.a.hearn@soton.ac.uk

Igor Filatotchev

King's Business School, King's College London, UK,
and Vienna University of Economics and Business, Austria
Email: igor.filatotchev@kcl.ac.uk

Marc Goergen*

IE Business School, IE University, Spain, and
European Corporate Governance Institute, Belgium
Email: marc.goergen@ie.edu

Final version published in *Journal of Corporate Finance* (2025), 102763,
<https://doi.org/10.1016/j.jcorpfin.2025.102763>

Abstract:

We explore differences in the levels of dispersed ownership that lead to a returns-based free float hedging factor in addition to size, which augments the capital asset pricing model (CAPM) in explaining the cross-section of stock returns. Using the S&P 1500 stocks in the US between 1985 and 2023, the results support the advantages of free float within a three-factor CAPM including size over alternative models based on liquidity, book-to-market value, and momentum. We argue that this yields a useful means for hedging effectively against the risks associated with the fundamental underlying likelihood of expropriation in a specific firm based on its ownership structure.

JEL classification: G11, G12, G15, O55

Keywords: CAPM; Free float, Ownership and control; investor protection

Acknowledgments: The authors are grateful for helpful comments received from participants at the 2023 annual meetings of the International Corporate Governance Society at IE Business School, Madrid, Spain. The research reported in this paper was partially funded by the Spanish State Research Agency (AEI) /10.13039/501100011033) – Grant No. PID2020-115982RB-C2 and Grant No. PID2023-148086NB-C22 granted by MCIU /AEI /10.13039/501100011033 / FEDER, UE.

* Corresponding author.

1. Introduction

The proportion of dispersed ownership, or free float, of firms is a visible outcome of firm-level protections provided for minority portfolio investors. Such protections are afforded through the interplay between external and internal governance factors (Helwege, Pirinsky, and Stulz, 2007; Hearn, Phylaktis, and Piesse, 2017). In a national context, the former are typified by the stringency of laws, stock exchange regulations, and listing requirements alongside their enforcement, as well as analyst coverage (Gompers, Ishii, and Metrick, 2003). The latter center on the bonding expenditures or credible commitments to uphold minority rights emanating from insiders and controlling shareholders within a firm (Charitou, Louca and Panayides, 2007). Together, the influences that shape firms' free float mirror the underlying motivation of insiders and controlling shareholders toward the expropriation of minority investors (e.g., Giannetti and Simonov, 2006; Bae, Baek, Kang, and Liu, 2012). This, in itself, is almost impossible to measure *ex ante* and is equally difficult to quantify *ex post* (Dyck and Zingales, 2004). In this paper, we ask whether variation in free float across a cross-section of stock returns is an overlooked factor that should be considered in asset pricing models.

The level of dispersed ownership in firms arises from a trade-off between two opposing perspectives. On the one hand, the dispersion of a firm's shareholding structure is associated with increased external financing to competitively facilitate capital expenditure and expansion (Helwege et al., 2007). On the other hand, an increasingly complex managerial task environment necessitates increased managerial discretion, which accompanies more concentrated control (Demsetz and Lehn, 1985). A firm's dispersed ownership arises from a tradeoff between opposing influences. Dispersion in ownership can be achieved in two ways: by increasing the number of issued shares or by diluting the concentrated control of insiders and controlling shareholders (Helwege et al., 2007). This overlap between the size of the listing and the dispersion of ownership, or free float, necessitates consideration of an analogous, size-adjusted free-float metric alongside the traditional free float measure. Furthermore, a significant benefit of focusing on free float is its capture of the aggregate commitment to minority property rights on the side of insiders and controlling shareholders. This circumvents thorny issues regarding the relative strength of influence of individual categories of shareholder and instead reflects the outcome of the inevitable discourse among them.

Our theoretical approach is based on Merton's (1973) intuition that minority portfolio investors' utility of wealth depends on how it can be used to generate future consumption and on the portfolio opportunities that will be available to move wealth through time for future consumption. Critically, wealth is contingent on stochastic state variables related to specific future consumption investment risks, such as the relative prices of consumption goods and the risk-return trade-offs in capital markets.

Moreover, changes in wealth over time are attributable to capital gains in equity assets held within portfolios. A drawback of Merton's (1973) approach is its complex mathematical exposition, which led to Fama (1996) and Fama and French (1996) outlining a simple, tractable intertemporal capital asset pricing model (ICAPM). This model is a multifactor efficient augmentation of the mean-variance efficient market factor of Sharpe's (1964) and Lintner's (1965) underlying CAPM, with additional returns-based factors mimicking state variables.

However, a consequence of the simplicity of this ICAPM derived from the addition of factors to the CAPM's market factor is that over the last 30 years, it has spawned considerable "data mining" for supplementary factors (Fama and French, 2018: 237). These factors are typically based on a multitude of financial ratios derived from firms' balance sheets. Moreover, following inclusion within an augmented CAPM, such factors often exhibit strong out-of-sample modeling performance (Fama and French, 2018) despite lacking theoretical support for their inclusion (Maio and Santa-Clara, 2012). We address this shortcoming by proposing free float as a factor, given its suitability as a proxy for a state variable related to the likelihood of expropriation of minority portfolio investors by insiders and controlling shareholders. The likelihood of this expropriation is a mirror reflection of the degree of protection of minority property rights provided by a firm (Albuquerque and Wang, 2008). Such firm-level commitments are essential in supporting and maintaining a dispersed ownership base for otherwise disenfranchised minority investors.

This leads to our first contribution to the literature based on our proposed free float factor as a proxy for a state variable associated with the likelihood of expropriation, given its potential far-reaching impact on all aspects of firm finances and performance. Our multifactor models containing a free-float factor mimicking portfolio, FMP, as a factor are generically consistent with Ross's (1976) APT and Merton's (1973) ICAPM. However, inspecting such consistency in detail by conducting appropriate empirical tests as undertaken in the case of the ICAPM by studies such as Maio and Santa-Clara (2012) or for the arbitrage pricing theory, or APT, framework by studies such as Cooper et al. (2021) is beyond the scope of our study and is left for future research.

Following Gompers et al. (2003), we argue that if the level of dispersed ownership matters for firm performance, then the stock price should quickly adjust to any relevant change in the firm's free float. Moreover, if stock prices fully reflect the appropriate discounting associated with free float, then the expected returns on the stock will be unaffected in the longer term after a short period of adjustment following any change in the free float. However, if free float matters but is not incorporated immediately into stock prices, then the realized returns on the stock would differ systematically from equivalent securities (Gompers et al., 2003: 121). Our theoretical approach is based on Giannetti and Koskinen's

(2010) assumption of mild segmentation within a given universe. This is defined by impediments to arbitrage trading that would otherwise close price differentials, leading to stock prices being determined locally by laws of supply and demand. Furthermore, the universe is defined by a finite number of listed firms alongside a finite pool of investors, the latter divided into minority portfolio investors and a combination of their insider and controlling shareholder counterparts. We assume that there is a continuum of variation in the likelihood of the expropriation of minority investors across firms, which is mirrored in firm-level protections provided for minority property rights (Giannetti and Koskinen, 2010), which, in turn, is reflected in the levels of free float.

We argue that higher likelihoods of expropriation are motivated by higher levels of concentrated ownership by insiders and controlling shareholders and lower protections provided for minority investors, as reflected by a lower free float. The reduction in free float and firm-level commitments to property rights protection leads to a reduction in demand by minority portfolio investors and a subsequent decrease in stock price (Giannetti and Koskinen, 2010), despite the potentially offsetting influence of reduced supply in the form of lower free float. The expropriation risk from holding the stocks of firms with a lower free float leads to higher expected returns to compensate minority investors for holding such stock. In contrast, lower likelihoods of expropriation are motivated by improved monitoring—either by external stakeholders or by minority investors (Giannetti and Koskinen, 2010)—rendering expropriation prohibitively costly. This, in turn, implies higher capital gains accruing universally to all shareholders, which then stimulates minority portfolio investor demand and subsequent stock price increases. The reduced risk of minority expropriation is reflected in lower expected returns. In summary, we implicitly consider a notional average level of firm-level protections of minority property rights—reflected in the level of free float and mirroring the underlying likelihood of expropriation. Our second contribution to the literature is proposing that minority investors expect a premium when they participate in firms with low free float (weaker protections against expropriation) and a discount when they participate in firms with high free float (stronger protections against expropriation).

To test the validity of the association between free float and a pricing premium we hypothesize, we undertake a “horse race” style comparison (e.g., Cooper and Maio, 2019a; Hou, Mo, Xue, and Zhang, 2019; Cooper, Ma, Maio, and Philip, 2021) contrasting eight augmented CAPM models, each drawing on a variety of returns-based factors in addition to the market factor. The first is the single-factor CAPM, which is contrasted with the second model, the Fama and French (1993) three-factor (FF3F) model, based on the CAPM and augmented by additional factors capturing cross-sectional differences in firm size and book-to-market value. The third model is the Carhart (1997) four-factor model, which augments

the preceding FF3F model with an additional momentum factor (thus, the Carhart 4F model). The fourth model is a two-factor model based on the CAPM and an additional illiquidity factor (the ILLIQ2F model), which captures differences in the cross-section of firms in relation to a multidimensional “trading speed” measure of liquidity developed by Liu (2006). Next, we consider free float as a valuation factor, henceforth *FFL*, which is a returns-based proxy for its variation across the cross-section of stock returns. Next, we augment the single-factor CAPM with the *FFL* factor to form the fifth model, a two-factor free-float model (the FFL2F model). Then, we consider a final set of three models all of which include the size-adjusted free-float factor, *FFL**, along with size. Hence, this sixth model is the traditional CAPM augmented by size and size-adjusted free float (the FFL3F model). The seventh model follows from this in augmenting the market excess returns of the CAPM and size with additional book-to-market value before also adding the *FFL** factor (forming the FFL4F model). Finally, the eighth model further augments this preceding model with momentum, leading to the full model comprising market excess returns, size, book-to-market value, momentum, and the *FFL** factors (the FFL5F model).

We focus on the US stock market with this universe comprising the stocks of the S&P 1500 index from January 1985 to December 2023. There is considerable diversity in corporate ownership (Anderson, Mansi, and Reeb, 2003; Anderson, Duru, and Reeb, 2009), reflected in marked variation in the free float. The results suggest that an unpriced premium is associated with the size-adjusted free float. A battery of tests is undertaken, with the first being time-series regression analysis applying four traditional asset pricing models, namely the CAPM, FF3F, Carhart 4F, and ILLIQ2F, on 10 decile portfolios following the cross-section of stock returns having been sorted by free float. The findings allude to an unexplained anomaly that is unpriced. We proceed as follows. First, we apply the Gibbons, Ross, and Shanklin (GRS; 1989) test statistics. Second, we apply the Cooper et al. (2021) two-step time-series cross-sectional procedure to form a “constrained” cross-sectional R-squared metric, R_C^2 , from the estimated pricing errors, or regression alphas, obtained from the first time-series step. Both the GRS and the R_C^2 facilitate the examination of each model as applied to groups of 10-decile portfolios sorted on 14 anomaly variables (e.g., Fama and French, 2008). The evidence reveals significant support for the relative strength of the FFL3F and FFL4F models. This statistical support for models including the size-adjusted *FFL** is continued following the application of the two-step time-series cross-sectional procedure and accompanying cross-sectional R-squared metric, R_C^2 , on the aggregate sample of 140 anomaly portfolios in total and 84 extreme portfolios as subsamples. Finally, we perform spanning regression tests for the seven models other than the single-factor CAPM; the tests reveal consistent

statistical support for the FFL3F and FFL4F models while further substantiating Fama and French’s (2018) claim of the redundancy of the “high minus low” (HML) book-to-market factor in asset pricing.

The paper proceeds as follows. Section 2 describes the sample selection and data sources. Sections 3 and 4 discuss the asset pricing methodology and then outline the techniques used to construct the FMPs, accompanied by summary statistics. Section 5 contains the preliminary analysis, while Section 6 reviews the results from the application of the various asset pricing tests. The final section concludes.

2. Data

2.1 Data

The final sample is the outcome of a series of screening stages, first at the market level, then in terms of data availability, and finally, at the individual stock level. All data are sourced from Refinitiv Datastream. We include all US firms that are constituents of the S&P 1500 benchmark index. Additionally, we disaggregate this sample into firms included in the large-cap S&P 500, mid-cap S&P 400, and small-cap S&P 600 indices, the three forming the aggregate S&P 1500 index. To mitigate potential survivorship bias, we additionally take the available updated constituent lists for the S&P 1500 index in December 1990, 1995, 2000, 2005, 2010, 2015, and 2023. This ensures consistent coverage throughout the sample period, ranging from January 1, 1985, to December 31, 2023.

Stocks are screened based on their inclusion in the above major blue-chip indices, where the lists are maintained within Datastream. Note that we exclude foreign firms. We use the stocks of blue-chip indices as they conform to international, and to a lesser extent, domestic *investability* requirements in terms of marketability and accessibility (minimal foreign ownership restrictions), while at the same time avoiding the thorny issue of imposed bias from pre-screening stocks based on predetermined minimum price criteria. This price pre-screening is evident in Hou, Karolyi, and Kho’s (2011) study, which focuses on a sample of 49 countries. Use of the blue-chip index stocks also facilitates conformity with asset diversification assumptions regarding asset market integration, which is essential to the CAPM, and thus avoids difficult issues regarding intra-market segmentation, which is particularly prevalent in very large stock markets.

The data include only single-class common stock and excludes preference stock, dual-class stock, warrants, convertibles, REITs, closed-end funds, exchange-traded funds, and depository receipts. Finally, following Ince and Porter (2003), any return above 300% that is reversed within one month is treated as *missing*; that is, if R_t or R_{t-1} is greater than 300%, and $(1 + R_t) \times (1 + R_{t-1}) - 1 < 50\%$, then R_t and R_{t-1} are set to *missing*. Following Hou et al. (2011), we ensure that the accounting ratios

are known before the returns and thus match the end-of-year financial statement data for year $t - 1$, with monthly returns from December of year t to December of year $t + 1$. We use the inverse of the market-to-book ratio to calculate the book-to-market value ratio, B/M (see Panel B of the appendix). In addition, size is defined as the market value or capitalization of equity, $MCAP$, at the end of December of year t , while momentum, MOM , for month t is the cumulative return from month $t - 12$ to month $t - 2$, skipping month $t - 1$ to avoid microstructure biases, such as bid-ask bounce or nonsynchronous trading (Roll, 1984). Our measure of illiquidity is the Liu (2006) multidimensional “trading speed” metric, LIU , which is the standardized turnover-adjusted number of zero daily trading volumes over the previous 12 months (see Panel A of the appendix for a detailed definition). This can be interpreted as the turnover-adjusted number of zero daily trading volumes over the previous year and is calculated at the end of each month for each individual stock based on daily data.

2.2. Free-float metric

Free float is defined as the mean monthly percentage of free-float shares, which is the proportion of firms’ total issued shares outside the control of block holders. Information on free float is easy to access for a wide cross-section of stocks. Refinitiv Datastream categorizes block holders into the following seven types: government, cross-shareholding networks, pension funds, investment companies, employees and family, other block entities, and foreign block holders (see Panel B of the appendix).

We find¹ that the US is dominated by institutional investor holdings, referred to in Datastream as “investment companies”, while free float systematically decreases from the average across firms that are in the large-cap indices (S&P 500) to their small-cap counterparts (S&P 600). Similarly, upon progression from large- to small-cap firms, there is a reduction in market capitalization and momentum, while the book-to-market value ratio, monthly returns, and Liu (2006) illiquidity metrics steadily increase.

2.3. Anomaly portfolios: Test assets

The previous asset pricing literature has identified hundreds of potential valuation factors based on anomaly characteristics explaining the cross-section of stock returns (see Fama and French, 2018; Hou, Xue, and Zhang, 2020). In this paper, we focus on 14 of the most prominent, including free float. Each factor is formed by sorting stocks into 10-decile portfolios based on each stock’s relative ranking in terms of the underlying anomaly variable, from the highest decile (D10) to the lowest (D1). Rebalancing

¹ See Table 1 in the online appendix.

occurs annually in December. A detailed outline of the construction of each and the sourcing of the data is provided in Panel B of the appendix. Below, we introduce the 14 variables, which form the right-hand side test assets in subsequent asset pricing tests.

The first variable is *Free Float*, which is defined as the percentage of shares available to minority investors at any given time, that are not closely held (i.e., it is the opposite of concentrated ownership). Next, $\beta^{pre-rank}$ (pre-ranking beta) is estimated using the preceding five years (rolling window) of previous monthly returns against the prevailing market benchmark index, the S&P 1500. A sizable number of studies, from Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973) to Frazzini and Pedersen (2014), find that the relation between the univariate market beta and the average stock return is flatter than predicted by Sharpe's (1964) and Lintner's (1965) CAPM. Fama and French (2016) qualify this beta anomaly as a purported violation of the CAPM. They advocate for the inclusion of a size variable *MCAP*, based on firms' market capitalizations, with stocks in the bottom and top size deciles referred to as microcaps and megacaps, respectively. The rationale for the *MCAP* factor is that smaller firms are more prone to earnings disparities during recessionary periods than their larger counterparts (Fama and French, 1993), although this has been questioned by Bebchuck, Cohen, and Ferrell (2009). *B/M* (book-to-market value) is defined as the ratio of the book value of equity to the market value of equity. Fama and French (1993) attribute persistent earnings variation to differences between value and growth stocks, as differentiated by their *B/M*, and this ratio has remained a cornerstone in asset pricing since then. However, more recent studies, such as Fama and French (2018), have questioned the importance of book-to-market value.

Next, we consider *Volatility*, defined as the variance in the daily closing price returns over the preceding 12 months. Ang, Hodrick, Xing, and Zhang (2006) find that stocks with highly volatile returns tend to have low average returns. Also included is *Asset Growth*, defined as the change in total assets from five years before the preceding year. Fama and French (2015a) attribute this change in total assets to investment or disinvestment by the firm in its own asset base. We also include *Accrual* (accruals), which is defined as the change in operating working capital per split-adjusted share from the fiscal year-end two years before the preceding year divided by the book equity per share in the preceding year. Our inclusion of accruals follows Sloan (1996) in attributing low returns to high accruals, while Fama and French (2018) argue that differences in accruals arise because accounting decisions cause book earnings to differ from cash earnings. Next, *OP* (operating profit) is defined as revenue minus the cost of goods sold, minus selling, general, and administrative expenses, minus interest expense, all divided by book equity in the preceding fiscal year. Fama and French (2015b) focus on operating profit as an anomaly that explains the variance in the cross-section of stock returns.

We also include six additional characteristics. *DY* (dividend yield) is defined as the total dividends paid out from July of the preceding year to June of the current year divided by the market equity at the end of June of the current year. Hou et al. (2021) argue that the dividend yield is directly related to variation in the cross-section of stock returns. *P/CF* (price to cash flow) is defined as the stock price to cash flow per share. It is measured as the market value of equity at the end of December of the previous year divided by the cash flows for the preceding fiscal year. *NSI* (net stock issues) is defined as the change in the natural logarithm of split-adjusted shares outstanding from two fiscal years prior to that of the preceding fiscal year. The inclusion of this variable follows from share repurchases tending to be followed by large average returns (Ikenberry, Lakonishok, and Vermaelen 1995), and average returns after share issues tending to be low (Loughran and Ritter, 1995). *P/E* (price-to-earnings ratio) is defined as the market value of equity at the end of December in the preceding year divided by earnings, which is defined as income before extraordinary items for the preceding fiscal year. The final two are *CAPX Growth* and *Sales Growth*. *CAPX Growth* is defined as the change in capital expenditure reported three years prior to that for the preceding year over the initial value three years prior. *Sales Growth* is defined in terms of current gross value of sales divided by that of six years prior.

3. Fama-MacBeth cross-sectional methods

The long-standing intuition (see Liu, Stambaugh, and Yuan, 2019) regarding value effects is that they scale a firm's equity price by equity characteristics, which are typically accounting balance sheet ratios, in relation to their impact on the expected returns. A scaled price is viewed as an effective proxy for the expected return; that is, a higher (lower) expected return implies a lower (higher) current price.

Our goal is to conduct a horse-race style comparison among equity characteristics first for the overall sample of stocks constituent to the S&P1500 universe and then for disaggregated big and small size subcategories of stocks drawn from within this universe. The horse-race style analysis involves comparison of eight Fama and MacBeth (1973) two-step cross-sectional regressions of individual monthly stock returns on the equity characteristics. The two steps comprise a first stage of regressions run cross-sectionally on a month-by-month basis across firms, followed by a second stage undertaking an averaging of these estimated monthly coefficients across the entire sample time frame. The choice of variables included in each Fama-MacBeth regression mirrors those underlying FMPs included in the later asset pricing models, considered in the next section.

The two size-related subsamples of main universe are formed from sorting constituent stocks into 10 size-portfolios based on their market capitalizations each December. The biggest subsample comprises stocks constituent to extreme three biggest size portfolios, while the smallest subsample

comprises stocks constituent to the extreme three smallest portfolios. Consideration of these subsamples facilitates consideration of consistency in relationships between stock returns and the variables across the extremities of the universe.

The first three Fama-MacBeth regressions mimic the CAPM, FF3F, and Carhart 4F models. They are outlined as follows:

$$R_{i,t+1} = \lambda_{i,0} + \lambda_{i,t}^{pre-rank} \beta_{i,t}^{pre-rank} + \epsilon_{i,t} \quad (1)$$

$$R_{i,t+1} = \lambda_{i,0} + \lambda_{i,t}^{pre-rank} \beta_{i,t}^{pre-rank} + \lambda_{i,t}^{MCAP} MCAP_{i,t} + \lambda_{i,t}^{B/M} B/M_{i,t} + \epsilon_{i,t} \quad (2)$$

$$R_{i,t+1} = \lambda_{i,0} + \lambda_{i,t}^{pre-rank} \beta_{i,t}^{pre-rank} + \lambda_{i,t}^{MCAP} MCAP_{i,t} + \lambda_{i,t}^{B/M} B/M_{i,t} + \lambda_{i,t}^{MOM} MOM_{i,t} + \epsilon_{i,t} \quad (3)$$

where in expression (1) $R_{i,t+1}$ is the individual asset's returns and $\beta_{i,t}^{pre-rank}$ is pre-ranking β . The corresponding second step focuses on the time-series average $\frac{1}{T} \sum_{t=1}^T \hat{\lambda}_{i,t}^{pre-rank}$ of the estimates of $\hat{\lambda}_{i,t}^{pre-rank}$, and it can be interpreted as the value effect for the pre-ranking β . In expression (2), in addition to the previously defined variable, $MCAP_{i,t}$ and $B/M_{i,t}$ are the logarithmically transformed market capitalization and the book-to-market value ratio, respectively. $\lambda_{i,t}^{MCAP}$ and $\lambda_{i,t}^{B/M}$ are the exposures attributable to these characteristics, at each time point. This leads to a second step focusing on the time-series averages $\frac{1}{T} \sum_{t=1}^T \hat{\lambda}_{i,t}^{pre-rank}$, $\frac{1}{T} \sum_{t=1}^T \hat{\lambda}_{i,t}^{MCAP}$, $\frac{1}{T} \sum_{t=1}^T \hat{\lambda}_{i,t}^{B/M}$ of the estimates of $\hat{\lambda}_{i,t}^{pre-rank}$, $\hat{\lambda}_{i,t}^{MCAP}$, $\hat{\lambda}_{i,t}^{B/M}$, which can be interpreted as the value effects. Correspondingly, in expression (3), in addition to the previously defined variables above, $MOM_{i,t}$ is the estimation coefficient of MOM from a preliminary time-series regression of asset returns on pre-ranking β , $MCAP$, B/M , and MOM . $\lambda_{i,t}^{MOM}$ is the exposure attributable to $MOM_{i,t}$, at each point in time. This in turn leads to a second step, which includes the additional time-series average $\frac{1}{T} \sum_{t=1}^T \hat{\lambda}_{i,t}^{MOM}$ of the estimate of $\hat{\lambda}_{i,t}^{MOM}$, which can be interpreted as the value effect.

Next, we estimate regression models based on two-factor configurations comprising the initial pre-ranking β plus Liu's (2006) trading speed measure of liquidity, LIU (mimicking the LIQ2F model), and then replace the latter variable with *Free Float* (mimicking the FF2F model). These are defined as follows:

$$R_{i,t+1} = \lambda_{i,0} + \lambda_{i,t}^{pre-rank} \beta_{i,t}^{pre-rank} + \lambda_{i,t}^{LIU} LIU_{i,t} + \epsilon_{i,t} \quad (4)$$

$$R_{i,t+1} = \lambda_{i,0} + \lambda_{i,t}^{pre-rank} \beta_{i,t}^{pre-rank} + \lambda_{i,t}^{Free Float} Free Float_{i,t} + \epsilon_{i,t} \quad (5)$$

where in both expressions (4) and (5), in addition to the variables previously defined, $LIU_{i,t}$ and $Free Float_{i,t}$ are the monthly values of Liu's (2006) liquidity metric and the percentage free float, respectively. Correspondingly, $\lambda_{i,t}^{LIU}$ and $\lambda_{i,t}^{Free Float}$ are the exposures attributable to each underlying equity characteristic, $LIU_{i,t}$ and $Free Float_{i,t}$, at a given time. The second step focuses on the time-series averages $\frac{1}{T}\sum_{t=1}^T \hat{\lambda}_{i,t}^{pre-rank}$, $\frac{1}{T}\sum_{t=1}^T \hat{\lambda}_{i,t}^{LIU}$, $\frac{1}{T}\sum_{t=1}^T \hat{\lambda}_{i,t}^{Free Float}$ of the estimates of $\hat{\lambda}_{i,t}^{pre-rank}$, $\hat{\lambda}_{i,t}^{LIU}$, and $\hat{\lambda}_{i,t}^{Free Float}$, which can be interpreted as the value effects.

Finally, we estimate our last three regressions, which mirror multifactor models based on free float. The first of these includes a pre-ranking β alongside $MCAP$ and $Free Float$ and is defined as:

$$R_{i,t+1} = \lambda_{i,0} + \lambda_{i,t}^{pre-rank} \beta_{i,t}^{pre-rank} + \lambda_{i,t}^{MCAP} MCAP_{i,t} + \lambda_{i,t}^{Free Float} Free Float_{i,t} + \epsilon_{i,t} \quad (6)$$

The second of these is based on the preceding expression (6) further including the book-to-market value, B/M , and defined as:

$$R_{i,t+1} = \lambda_{i,0} + \lambda_{i,t}^{pre-rank} \beta_{i,t}^{pre-rank} + \lambda_{i,t}^{MCAP} MCAP_{i,t} + \lambda_{i,t}^{B/M} B/M_{i,t} + \lambda_{i,t}^{Free Float} Free Float_{i,t} + \epsilon_{i,t} \quad (7)$$

The final regression is based on expression (7) with additional momentum and is defined as:

$$R_{i,t+1} = \lambda_{i,0} + \lambda_{i,t}^{pre-rank} \beta_{i,t}^{pre-rank} + \lambda_{i,t}^{MCAP} MCAP_{i,t} + \lambda_{i,t}^{B/M} B/M_{i,t} + \lambda_{i,t}^{MOM} MOM_{i,t} + \lambda_{i,t}^{Free Float} Free Float_{i,t} + \epsilon_{i,t} \quad (8)$$

where all the variables and coefficients are as previously defined.

4. Asset pricing test methodology

4.1 Asset pricing models

Our analysis is based on eight asset pricing models, all utilizing time-series ordinary least squares regressions ubiquitous to the asset pricing literature. The first is the traditional CAPM, defined as follows:

$$R_{it} - r_{f,t} = \alpha_{i,t} + \beta_{i,M}(R_{M,t} - r_{f,t}) + \epsilon_{i,t} \quad (9)$$

where $R_{it} - r_{f,t}$ are the returns of the portfolio or test asset i ($i = 1, \dots, N$) in excess of the risk-free rate ($t = 1, \dots, T$), $\alpha_{i,t}$ is the Jensen alpha or a $1 \times T$ vector of constant coefficients, and $R_{M,t} - r_{f,t}$ is the

value-weighted market return in excess of the risk-free rate of return (i.e., the one-month US Treasury rate). $\beta_{i,M}$ is a vector of factor loadings for asset i on the market excess returns, and $\varepsilon_{i,t}$ represents the idiosyncratic i.i.d. errors, which are allowed to have a limited correlation among returns.

Fama and French (1993) extended this basic specification to additionally capture the relation between the average return and the size and the accounting book-to-market value, B/M , leading to a three-factor model:

$$R_{it} - r_{f,t} = \alpha_{i,t} + \beta_{i,M}(R_{M,t} - r_{f,t}) + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \varepsilon_{i,t} \quad (10)$$

where SMB_t is the difference between the returns on a diversified portfolio of small stocks minus those of an equally diversified portfolio of big stocks, and HML_t is similarly the difference between the returns on a diversified portfolio of high B/M stocks minus those of a diversified portfolio of low B/M stocks. The formation of SMB_t and HML_t uses a 3×3 double-sort procedure in which stocks are sorted into five-quintile portfolios based on their size, or market capitalization, each of which is further sorted into another five-quintile portfolios based on B/M . At any time, stocks with missing values for either characteristic are omitted, as are stocks with negative book-to-market values. FMPs related to size are created from the average returns on small portfolios minus those on large portfolios (based on the SMB_t factor) and similarly with high book-to-market value portfolios minus low book-to-market portfolios (based on the HML_t factor). Portfolio rebalancing takes place annually in December. The SMB_t and HML_t factors are formed from value-weighted returns. $\beta_{i,SMB}$ and $\beta_{i,HML}$ are the vectors of factor loadings for asset i on the SMB_t and HML_t factors, respectively.

Carhart (1997) augmented the Fama-French three-factor model with a fourth factor related to momentum or persistence in stock price returns. This is defined as follows:

$$R_{it} - r_{f,t} = \alpha_{i,t} + \beta_{i,M}(R_{M,t} - r_{f,t}) + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \beta_{i,UMD}UMD_t + \varepsilon_{i,t} \quad (11)$$

where UMD_t is the difference between the returns on a diversified portfolio of high-performing or “up” stocks minus those of an equally diversified portfolio of underperforming “down” stocks. This involves sorting stocks into 10-decile portfolios based on momentum, defined as the cumulative return over the preceding 12 months. The FMP for momentum follows the Jegadeesh and Titman (1993) 12-month/12-month strategy, whereby the monthly returns are the average of 12 individual strategies of buying the winning, or “up,” decile portfolio and selling the losing, or “down” decile portfolio. Rebalancing occurs

monthly². To minimize the influence of potential microstructural biases such as bid-ask bounce effect, we construct the momentum UMD_t FMP for month t , which is the cumulative return from month $t - 12$ to month $t - 2$, skipping month $t - 1$. This FMP is also formed from value-weighted returns. $\beta_{i,UMD}$ is the vector of factor loadings for asset i on the UMD_t factor.

Liu (2006) proposed augmentation of the traditional CAPM with an additional liquidity factor. Liu's measure is based on the turnover-adjusted number of zero daily volume, which better captures the multidimensional nature of liquidity and overcomes shortcomings in numerous unidimensional measures prevalent in the literature. This two-factor model is defined as follows:

$$R_{it} - r_{f,t} = \alpha_{i,t} + \beta_{i,M}(R_{M,t} - r_{f,t}) + \beta_{i,ILLIQ}ILLIQ_t + \varepsilon_{i,t} \quad (12)$$

where $ILLIQ_t$ is the difference between the returns on a diversified portfolio of highly illiquid stocks minus those of an equally diversified portfolio of low illiquidity stocks. Stocks are sorted into 10-decile portfolios, and the FMP is formed from the returns difference between high and low illiquidity decile portfolios. The FMP is based on annual rebalancing each December, as in Liu (2006), and is value weighted to minimize the effects of illiquid microcap stocks. $\beta_{i,ILLIQ}$ is a vector of factor loadings for asset i on the $ILLIQ_t$ factor.

We propose a two-factor model based on the traditional CAPM augmented by a factor that captures the relation between average returns and free float, FFL , defined as follows:

$$R_{it} - r_{f,t} = \alpha_{i,t} + \beta_{i,M}(R_{M,t} - r_{f,t}) + \beta_{i,FFL}FFL_t + \varepsilon_{i,t} \quad (13)$$

where FFL_t is the difference between the returns on a diversified portfolio of stocks with a low percentage free float minus those of an equally diversified portfolio of high-percentage free float stocks. Stocks are sorted into 10-decile portfolios, and the FMP is formed from the returns difference between low and high free-float decile portfolios. The FMP is based on annual rebalancing each December and is value weighted again to minimize the effects of illiquid microcap stocks. $\beta_{i,FFL}$ is the vector of factor loadings for asset i on the FFL_t factor.

Furthermore, we define the size-adjusted free float, FFL_t^* , as the low-minus-high difference in returns following a second stage of portfolio sorting based on free float after an initial first stage of sorting based on size alone within a 3×3 double-sort procedure. We propose a three-factor model

² That is, the momentum FMP return for January 2001 is 1/12 of the return spread between the winners and losers from January to November 2000, 1/12 of the return spread between winners and losers from December 1999 to October 2000, etc.

comparable to the Fama-French three-factor model in expression (2) by replacing HML with the size-adjusted free-float factor, FFL_t^* :

$$R_{it} - r_{f,t} = \alpha_{i,t} + \beta_{i,M}(R_{M,t} - r_{f,t}) + \beta_{i,SMB}SMB_t + \beta_{i,FFL^*}FFL_t^* + \varepsilon_{i,t} \quad (14)$$

where β_{i,FFL^*} is the vector of factor loadings for asset i on the FFL_t^* factor. Next, we propose a four-factor model directly augmenting the Fama-French three-factor model in expression (10) augmented by the additional size-adjusted free-float factor, FFL_t^* :

$$R_{it} - r_{f,t} = \alpha_{i,t} + \beta_{i,M}(R_{M,t} - r_{f,t}) + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \beta_{i,FFL^*}FFL_t^* + \varepsilon_{i,t} \quad (15)$$

where β_{i,FFL^*} is a vector of factor loadings for asset i on the FFL_t^* factor. SMB_t is formed from the average of the two size factors formed from the two respective initial stage size sorts, first to form the HML_t factor and then subsequently to form the FFL_t^* factor. Finally, we consider the potential implications of momentum in augmenting the four-factor Carhart (1997) model in expression (11) with the additional size-adjusted free-float factor, FFL_t^* :

$$R_{it} - r_{f,t} = \alpha_{i,t} + \beta_{i,M}(R_{M,t} - r_{f,t}) + \beta_{i,SMB}SMB_t + \beta_{i,HML}HML_t + \beta_{i,UMD}UMD_t + \beta_{i,FFL^*}FFL_t^* + \varepsilon_{i,t} \quad (16)$$

We adopt value-weighted returns across all the FMPs and test portfolios in the subsequent analysis. This is intended to mitigate the effects of the proliferation of microcap stocks across all factor-sorted portfolios on which the FMP formation is based. This issue is especially important, given that our breakpoints are freely determined by an even distribution of stock numbers based on the sorting procedure of the underlying variable of interest. This procedure is especially susceptible to the spread of microcaps throughout the factor-sorted portfolios. In particular, microcaps can inflate the magnitude of anomalies, especially when combined with equal-weighted returns. Fama and French (2008) highlight that microcaps account for at most 3% of the aggregate market capitalization of the NYSE-Amex-NASDAQ universe, from which the S&P 1500 index constituents are drawn, but account for about 60% of the total number of stocks.

4.2 Time-series regressions and differentiation between models

4.2.1 Time-series regressions

We follow the time-series approach of Black et al. (1972) and studies, such as Fama and French (1993, 1996, 2015b) and Hou, Xue, and Zhang (2015). We use 11 test portfolios. These comprise 10 portfolios formed from sorting the cross-section of stock returns by each firm's percentage free float plus a "high-minus-low" hedging portfolio from the difference in returns from the highest value decile minus those of the lowest free float decile. The test portfolio excess returns are formed from the value-weighted average return on each of the 11 portfolios minus the risk-free rate. We regress the excess returns for the 11 test portfolios on those of the value-weighted market portfolio plus the returns of the FMPs for each of the four asset pricing models outlined in expressions (9) to (12). The time-series slopes are interpreted as factor loadings that inform how various combinations of these FMPs explain the average returns across portfolios. Our objective with this exercise is to ascertain the extent to which any of the traditional asset pricing models, namely CAPM, FF3F, Carhart 4F, and ILLIQ2F, between expressions (9) to (12), are able to explain the free-float anomaly within the cross-section of stock returns.

A large and significant Jensen alpha is viewed as an *abnormal return* that cannot be attributed to any of the included FMPs or "the return in excess of what could have been achieved by passive investments in any of the factors" (Gompers et al., 2003: 122). Importantly, the focus of this exercise is to ascertain whether any of the initial four asset pricing models are effective in explaining the free-float anomaly in the cross-section of stock returns. Therefore, if the Jensen alphas are consistently large and statistically significant, this implies a weakness in traditional asset pricing models in their ability to capture the free-float anomaly in the cross-section of stock returns. This would theoretically justify the potential inclusion of free float within an asset pricing framework.

4.2.2 Differentiation between models

We undertake further differentiation between all eight asset pricing models, in expressions (9) to (16), using 11 test portfolios based on each of 14 anomaly variables (e.g., free float, dividend yield, *NSI*, and accruals). These comprise 10 portfolios formed from sorting the cross-section of stock returns by each of the 14 respective anomalies, and then an eleventh hedging "high-minus-low" portfolio from the returns difference between the highest and lowest decile portfolio returns. In each of the 14 anomalies, the 11 test assets are expressed in excess returns from the average test portfolio returns minus risk-free rate.

Differentiation between the eight asset pricing models is undertaken by the analysis of six sensitivity statistics. The first four of these originate from the application of the eight asset pricing

models to the 14 anomaly groups of 10 test portfolio returns. The fifth and sixth sensitivity statistics then relate to each of the eight asset pricing models applied to the 14 respective high-minus-low anomaly hedging portfolios.

The first sensitivity statistic is the average of the R^2 statistics obtained from the application of each of the eight asset pricing models in turn as applied to the 10-test portfolio returns within each of the 14 respective anomaly groups. The second is the “constrained” cross-sectional R-squared, R_C^2 , proposed in Maio and Santa-Clara (2017) and Cooper and Maio (2019b). This statistic arises from considering that the time-series models in expressions (9) to (16) can also be expressed in the basic general form of

$$R_{it} - r_{f,t} = \alpha_{i,t} + \beta_i F_t + \varepsilon_{i,t} \quad (17)$$

where $R_{it} - r_{f,t}$ are the returns of portfolio or test asset i ($i = 1, \dots, N$) in excess of the risk-free rate ($t = 1, \dots, T$), $\alpha_{i,t}$ is the Jensen alpha or a $1 \times T$ vector of constant coefficients, F_t is a K -dimensional vector of common FMPs (returns-based factors) at t , with K being the number of FMPs included in the model, β_i is a K -dimensional vector of factor loadings for the excess return on asset i , and $\varepsilon_{i,t}$ represents the idiosyncratic i.i.d. errors, which are allowed to have a limited correlation among the returns. Next, we follow Maio (2017), Cochrane (2005), and Lewellen, Nagel, and Shanken (2010) in performing a two-step analysis. This analysis builds on an initial time-series regression step, which is outlined in expressions (9) to (16). Using expression (17) as an example, we form the “constrained” model for each of the test asset regressions across all 10-decile portfolios of each of the 14 anomalies:

$$\overline{R_i - r_f} = \hat{\beta}_{i,F} \bar{F} + \hat{\alpha}_{i,C} \quad (18)$$

and with an example again in the form of the FFL4F (from expression (15)):

$$\begin{aligned} \overline{R_i - r_f} = & \hat{\beta}_{i,M} (\overline{R_M - r_f}) + \hat{\beta}_{i,SMB} \overline{SMB} + \hat{\beta}_{i,HML} \overline{HML} \\ & + \hat{\beta}_{i,FFL^*} \overline{FFL^*} + \hat{\alpha}_{i,C} \end{aligned} \quad (19)$$

where $\overline{R_i - r_f}$ is the average of the time-series excess returns on test asset i ; $\overline{R_M - r_f}$, $\overline{SMB}$, $\overline{HML}$, and $\overline{FFL^*}$ are the averages of the time-series FMP returns; $\hat{\beta}_{i,M}$, $\hat{\beta}_{i,SMB}$, $\hat{\beta}_{i,HML}$, and $\hat{\beta}_{i,FFL^*}$ are the estimated factor loadings on each FMP obtained from the time-series regression; and $\hat{\alpha}_{i,C}$ is the pricing error.

Expressions (18) and (19) are considered constrained, as they restrict the risk price estimates to the factor means (see Maio and Santa-Clara, 2017). Therefore, rather than estimating an additional cross-sectional regression based on the FMP betas to obtain risk price estimates, we conduct a sensitivity

analysis based on expression (18) obtained from the application of each of the eight asset pricing models for each of the 10-decile portfolios representative of each of the 14 anomalies. This amounts to 10 test asset portfolios in each of the 14 separate cases. We assume $\hat{\alpha} = (\hat{\alpha}_{1,C}, \dots, \hat{\alpha}_{N,C})$ denotes the vector of alphas estimated from the first-stage time-series regression, which represents the pricing error in the second-stage cross-sectional constrained model. Then the R_C^2 proposed in Maio and Santa-Clara (2017) and Cooper and Maio (2019b) is defined as follows:

$$R_C^2 = 1 - \frac{Var_N(\hat{\alpha}_{i,C})}{Var_N(\bar{R}_i - r_f)} \quad (20)$$

where $\hat{\alpha}_{i,C}$ is the pricing error, $\bar{R}_i - r_f$ is the predicted mean of excess returns of the test asset (as calculated from the estimated beta coefficients) from expression (18), and Var_N is the variance.

Additionally, we follow Cooper and Maio (2019b) and Cooper et al. (2021) in undertaking a statistical bootstrap simulation exercise to assess the statistical significance of the second sensitivity statistic, R_C^2 , through the computation of p-values. These correspond to the proportions of artificial samples in which the pseudo statistics are higher than the corresponding sample estimates. In the simulation, we impose the condition that the factors are independent from the test portfolio returns (akin to the “useless factors” in Kan and Zhang, 1999), while preserving the correlations among factors in a given asset pricing model (i.e., expressions (9) to (16)). A fully detailed outline of the bootstrap simulation is presented in the online appendix.

The third sensitivity statistic is the mean absolute alpha (MAA) from the application of each asset pricing model across each set of 10-decile portfolios per each of the 14 anomalies in succession. The MAA is defined as follows:

$$MAA = \frac{1}{N} \sum_{i=1}^N |\hat{\alpha}_i| \quad (21)$$

where $\hat{\alpha}_i$ with $i = 1, \dots, N$ represents the first N moments, i.e., the pricing errors associated with the N testing assets. Note that we estimate all regression slopes as constants; therefore, time variation in the slopes is a potential problem. Similar to most asset pricing literature, our models and tests also assume there are no market frictions, such as transaction costs and taxes.

At this stage, it is worth noting that the R_C^2 metric overcomes shortcomings in the MAA (expression (21)) relating to the latter’s inability to relate the magnitudes of the pricing errors to the magnitudes of the raw portfolio risk premiums that we seek to explain in the first place. This is exemplified by the given model, possibly producing an average pricing error that appears to be large but

is actually small in comparison with the scale of the raw risk premiums that we are trying to explain. This is especially important in our case, as we have joint asset pricing tests involving many different anomalies and, thus, different magnitudes of risk premiums.

The fourth sensitivity statistic is the GRS statistic to test the validity of the null hypothesis that if an asset pricing model captures expected returns, then the intercept is indistinguishable from zero in the time-series regression of any asset's excess return (its return in excess of the risk-free rate) on the model's factor returns. The GRS statistic tests the validity of this hypothesis for each of the seven asset pricing models applied to 10-decile portfolios sorted based on the 14 anomalies (e.g., free float, dividend yield, *NSI*, and accruals) outlined in the preceding section. Therefore, for each of the 14 anomaly variables, we utilize 10 sorted decile portfolios as left-hand side test assets to which we apply the eight asset pricing models in succession.

The fifth and sixth sensitivity statistics are the time-series regression constant and associated t-statistic estimated for each of the hedging portfolios, namely high-minus-low, for each of the respective 14 anomalies and for each of the eight asset pricing models in turn. Here, the objective is to gauge if the premium associated with each of these 14 anomalies (high-minus-low spread return) is explained by any of the eight factor models. Our particular focus is in differentiating between those models including the *FFL* and its size-adjusted counterpart, *FFL**, and those excluding these free-float factors in terms of relative size and statistical significance of the regression alphas.

4.3 Joint time-series and cross-sectional methods

Here, we focus on two sensitivity statistics. The first is the mean absolute error or alpha obtained from the $\hat{\alpha}_{i,C}$ test assets, as defined in expression (18), which is estimated from second-stage cross-sectional regressions, following the first time-series stage, as outlined in expression (17). This is estimated for each of the eight asset pricing models in expressions (9) to (16). The second is the “constrained” cross-sectional R-squared, R_C^2 , as outlined in the preceding section. This is also accompanied by p-values arising from a statistical bootstrap simulation exercise, which provides a measure of the statistical significance of the R_C^2 . Both metrics are applied, as per each of the eight asset pricing models, across two groups of test assets. The first group, referred to as the overall sample, is comprised of 10 portfolios formed from a decile-sort of the 14 respective anomalies, which leads to a total of 140 test asset portfolios. The second group, referred to as extreme deciles, is comprised of the three highest and three lowest of the 10 portfolios formed from each of the respective sorts on the 14 anomalies. This leads to 6×14 , or 84 test asset portfolios.

4.4 Spanning regression tests

We utilize spanning regression tests as an informative and concise way to compare asset pricing models. Following Fama and French (2015b, 2018), Hou et al. (2020), and Barillas and Shanken (2017, 2018), spanning regression tests provide a useful means of comparing the efficacy of individual factors within a given model with the joint ability of all the other factors within the model to explain the variation in any given factor. Barillas and Shanken (2017, 2018) argue that for models with traded factors, the extent to which the combination of all other factors within a model is able to price the focal factor is all that matters for model comparison. Thus, if an individual factor is “spanned” by the other factors, then it is effectively redundant in asset pricing and in explaining the cross-section of stock returns in comparison to the other factors within the model.

Our starting point is estimating the maximum Sharpe ratio for each of the seven asset pricing models. Following MacKinlay (1995), we define a given model’s maximum Sharpe ratio as follows:

$$Sh(F)^2 = \bar{\mu}_F \Pi^{-1} \bar{\mu}_F' \quad (22)$$

where $\bar{\mu}_F$ is the vector of returns-based FMP means, including the mean of the excess returns on the initial market factor, and Π is the variance-covariance matrix of the FMP means. Second, following Fama and French (2018), we focus on the contribution of each FMP to the maximum Sharpe ratio associated with the overall asset pricing model. Fama and French define the incremental increase in the maximum Sharpe ratio arising from FMP_j over and above that attributable to the other FMPs collectively as α_j^2 / σ_j^2 . This expression implies that an FMP’s marginal contribution to a model’s maximum squared Sharpe ratio is small if the FMP’s expected return is explained well by other FMPs (α_i is close to zero), and its variation not explained by other factors (σ_j) is large, or both.

5. Preliminary analysis

5.1. Value effects

Table 1 reports the average slopes from the month-by-month Fama-MacBeth regressions applied to the overall sample (Panel A), followed by the small (Panel B) and large (Panel C) subsamples. Notably, the statistical significance of the pre-ranking β is generally negligible and minimal at best across all models and all three panels. This statistical significance becomes negligible after the inclusion of the *LIU* and *Free Float* equity characteristics within the last three models across the overall sample and the subsamples. The addition of *MCAP* is consistently negative and highly statistically significant across

all relevant models in the overall sample. However, this association becomes consistently positive across both Panels B and C. This association generally lacks statistical significance in the smaller subsample (Panel B), while having high statistical significance in the larger subsample (Panel C). This implies a possible size-related effect in the cross-section of stock returns.

Next, the addition of *B/M* is associated with a consistently negative coefficient, which is also highly statistically significant across all models that include this variable and across all three panels. The addition of *MOM* results in a positive and highly statistically significant coefficient, alluding to a momentum effect in the valuation of the stock returns. This evidence is consistent with Fama and French's previous (1992) findings and, more recently, Liu et al.'s (2019) findings. The addition of *LIU* similarly leads to a consistently negative coefficient across all models, which is notably highly statistically significant in the overall sample and the smallest size subsample (Panel B), while lacking significance in the largest-size subsample (Panel C). This evidence is suggestive of a potential liquidity effect, which is especially pertinent in smaller-sized stocks. The addition of *Free Float* leads to a consistently negative coefficient across all models and across all three samples. However, while this is highly statistically significant in the overall sample (Panel A), it is somewhat reduced in the two size-related subsamples. This is indicative of a potential free float-related effect.

Table 1 here

5.2. Descriptive statistics of factor mimicking portfolios

The means, standard deviations, autocorrelations, and cross-correlations of the monthly returns of the FMPs are reported in Table 2. Several observations are apparent. The first in Panel A is that the largest average monthly returns are also those with the highest statistical significance. The average excess returns of the market portfolio are 0.87% (t-statistic: 2.29), while those of the size (*SMB*) and book-to-market value (*HML*) FMPs are 1.30% (t-statistic: 8.04), and −0.15% (t-statistic: −0.86), respectively. The small size of the *HML* returns and their lack of statistical significance support empirical evidence from Fama and French (2015b, 2018) emphasizing the redundancy of *HML* in asset pricing.

The momentum (*UMD*) and liquidity (*ILLIQ*) FMPs have large and significant average monthly returns, with 0.84% (t-statistic: 3.80) and −1.59% (t-statistic: −4.77), respectively. These returns underscore the importance of momentum and liquidity. Finally, the average monthly returns attributable to the *FFL* FMP are visibly smaller than those of all the other factors at 0.21% (t-statistic: 1.62). The monthly average returns of the size-adjusted *FFL** are only a fraction of the size of all the other factors and are not significant (t-statistic: 0.57). Together, the evidence indicates strong support for potential

premiums associated with the market factor, as well as the *SMB*, *UMD*, *ILLIQ*, and *FFL* FMPs. In contrast, there is very little support for the *HML* FMP.

The analysis in Panel B of the correlations reveals negligible correlations among factors. Finally, in Panel C, concerns regarding autocorrelation are mitigated due to the low absolute size and minimal statistical significance over the 1-, 6-, and 12-month lagged periods.

Table 2 here

5.3. Descriptive statistics of free-float portfolios

Descriptive statistics for each of the 10 value-weighted free-float sorted portfolios are reported in Table 3. The evidence from Panel A regarding the distribution of stocks across the S&P 500, S&P 400, and S&P 600 indices reveals two distinct opposing transitions. There is a progressively increasing number of S&P 500 (large cap) stocks in the higher free-float portfolios, with the highest concentration in D10. The opposite trend is visible for S&P 600 (small cap) stocks, whose concentration increases in the lowest free-float portfolios with their highest concentration being in D1.

A comparison of the trading and descriptive statistics for the stocks sorted into the 10 free float value-weighted decile portfolios is provided in Panel B of Table 3. Five observations are apparent. First, there is evidence of a statistically significant difference in the average value- and equal-weighted monthly returns between D1 and D10. This is also accompanied by a progressive rise in the average monthly returns from D10 to D1. This implies statistical support for a free-float return-based premium.

Second, following Fama and French (2016) and Hou et al. (2020), we focus on the variation in the 14 anomaly variables associated with constituent firms within the 10 free-float portfolios. There is a visible increase in both *Free Float* and *B/M* from D1 to D10, while the opposite is true for *MOM*. All three variables have highly statistically significant ($p \leq 0.01$) differences in means between D1 and D10. The important takeaway is that there are significant, yet opposite trends in the percentage *Free Float* and *B/M* ratio, as well as momentum across the cross-section of stocks sorted by their free float.

Third, there are consistently statistically significant ($p \leq 0.01$) differences between the extreme lowest (D1) and highest (D10) *FFL* decile portfolios for all trading variables. *MCAP*, *Volume*, and *Price* (stock price) all increase progressively from D1 to D10, and the differences are highly statistically significant ($p \leq 0.01$). This is in line with a progressive increase (decrease) in liquidity (illiquidity) in the gradual reduction of *LIU* ($p \leq 0.10$). It is notable that some outliers disrupt the smoothness of this transition in the values for deciles D2 and D9.

Fourth, focusing on the ascending progression from the lowest to highest free-float decile portfolio, there is a parallel transition from low to high value for operating profit, *OP*, with the difference

between the extreme portfolios being statistically significant ($p \leq 0.01$). In contrast, there is the opposite transition, with the anomalies relating to *NSI* ($p \leq 0.10$), *Volatility* ($p \leq 0.01$), and *Asset Growth* ($p \leq 0.01$). These are all highest in D1 and decrease progressively toward D10. However, there is a notable lack of any progressive trend or statistically significant difference in terms of beta (β) and *Accrual*. Together, these findings imply that firms with more concentrated ownership are more illiquid, reflected in being prone to erratic price adjustments associated with volatility, while making more equity issues and engaging in higher investment, reflected in elevated asset growth.

Our final observation relates the dispersion of different types of block ownership over the 10 free-float portfolios as reported in Panel C of Table 3. Five of the seven categories of block ownership are highest in the lowest free-float decile (D1) and decrease progressively to negligible values in the highest free-float decile (D10). Notably, the single largest category of ownership is that of investment companies, which is close to or over 8% in all 10 decile portfolios.

Table 3 here

6. Results

The first series of time-series tests focuses on how well all existing non-FFL asset pricing models explain the time-series average returns of 10 decile-sorted free-float portfolios, ranging from lowest (D1) to highest (D10). Also included is a free-float hedging portfolio formed from the returns difference between extreme portfolios D1 and D10 (or low-minus-high). The results are reported in Table 4.

This exercise reveals substantial evidence regarding the relative weakness of alternative asset pricing models, namely the CAPM (Panel A), FF3F (Panel B), Carhart 4F (Panel C), and ILLIQ2F (Panel D), in explaining free float as an anomaly in the cross-section of stock returns. The first observation relates to the explanatory power or adjusted- R^2 increasing from its lowest values following the application of the single-factor CAPM (Panel A) to the next highest values following the application of the ILLIQ2F model (Panel D), with the highest values being attributable to both the FF3F (Panel B) and Carhart 4F (Panel C) models. This evidence is suggestive of the relative strength of the FF3F and Carhart 4F models in contrast to the CAPM and ILLIQ2F models.

Next, the regression constants or alphas are generally large and statistically significant in relation to the application of each of the four asset pricing models to the 10 free-float portfolio returns, as visible across Panels A to D. Similarly, the regression alphas are statistically significant on the free-float hedging portfolio (formed from low-minus-high returns) in the final column for all four models across Panels A to D. Together, the findings suggest that the four traditional asset pricing models are relatively weak in explaining the free-float anomaly within the cross-section of stock returns. This evidence not

only suggests the presence of a hitherto unpriced anomaly related to concentrated versus dispersed ownership but also the potential importance and justification for further exploration of a free-float valuation factor. This forms the focus of the remainder of this section. Moreover, similar evidence is visible from the analysis conducted on equal-weighted portfolios³, which supports our findings.

Table 4 here

Table 5 reports six further statistical tests. The first four originate from the sequential application of each of the eight asset pricing models to 14 groups of 10 portfolios. These have been formed by sorting the cross-section of stock returns on each of the 14 anomaly variables. The last two statistical tests relate to the regression constant, or alpha, and its accompanying t-statistic derived from the application of each of the eight asset pricing models on the hedging or difference portfolios for each of the 14 respective anomalies.

We start by focusing on the first four statistical tests. The first of these is the average R^2 for each of the eight respective models applied to each group of 10 portfolios formed through the sorting of the cross-section by 14 anomalies. In almost all of the 14 anomalies the average R^2 is highest, typically by approximately one percentage point, in the two largest multifactor models containing the size-adjusted *FFL**, namely FFL4F and FFL5F. These notably contain the size-adjusted *FFL** derived from the 3×3 double-sort process on size first and then free float. The second is the cross-sectional constrained R^2 (R_C^2), which is highest or least negative for the FFL3F model with it also being highest in a further two anomalies for the FFL4F model for half of the anomalies (i.e., 7 out of 14 cases). This evidence is also mirrored in the R_C^2 values of these two models' also having markedly lower p-values indicating statistical significance too. An important attribute of the R_C^2 statistic is that it can assume negative values, as is evident from several of the anomalies. Cooper and Maio (2019b: 1986) argue that such a negative R_C^2 “means that the model does worse than a simple cross-sectional regression containing just a constant.”

The third statistical test is that of the MAA, which is lowest in eight anomalies (8 of the 14) for the FFL3F model, with this being lowest for a further two cases regarding the application of the FFL4F model. The fourth statistical test is that of the GRS test statistic. This visibly is associated with an overwhelming rejection of the null hypothesis ($p \leq 0.005$) that all regression intercepts are jointly equal to zero across all eight models and all 14 anomalies, i.e., all 140 test portfolios. This largely uniform rejection of the GRS test statistic is in line with previous literature, with Fama and French (1993) for US stocks and with Hou et al. (2011) for a multi-country sample comprised of 49 major equity markets

³ The results are available in Table 1 in the online appendix.

worldwide. Moreover, the GRS test statistic is lowest for six of the anomalies (6 out of 14 cases) for FFL3F, while it is least in a further three cases for FFL4F. Together, this evidence provides compelling statistical support for the relative strength of the FFL3F and FFL4F models, which include the size-adjusted *FFL** factor derived from a 3×3 underlying size followed by free-float double sort.

The fifth and sixth statistical tests reveal that in half of the anomaly hedging portfolios, the FFL3F model has the smallest absolute-sized regression alphas and with least (negligible) statistical significance, underscoring its superiority among the eight asset pricing models. Furthermore, the next strongest model is the FFL4F, which leads to the smallest absolute-sized regression alphas, lacking in statistical significance for a further three anomalies. This evidence substantiates the importance of the combination of size and size-adjusted *FFL** within asset pricing frameworks. Similar evidence is visible from the analysis conducted on equal-weighted portfolios⁴, which supports our findings.

Table 5 here

Table 6 reports the results from the joint time-series analysis of the overall sample, comprising 14 anomaly sorts of 10-deciles (i.e., 140 test portfolios), and a smaller extreme subsample, comprising the three highest and three lowest decile portfolios (i.e., 84 test portfolios). The MAA is smallest for the FFL3F model across the overall sample (0.181) and the extreme subsample (0.221), implying its superiority over rival models. These findings are further corroborated by the evidence from R_C^2 , which supports the FFL3F model across both the overall sample ($R_C^2 = 0.539$, $p = 0.00$) and the extreme subsample ($R_C^2 = 0.531$, $p = 0.00$). A further notable observation is the R_c^2 statistic, which is negative in the LIQ2F model ($R_c^2 = -0.047$) for the extreme subsample, while being positive but very small in value (0.022) for the overall sample. This evidence of the redundancy of the LIQ2F model is further reinforced by the R_C^2 lacking in statistical significance at the 10% confidence margin (p-value of 0.44). In summary, at this stage, the evidence points to the superiority of the FFL3F model. Similar evidence is visible from the analysis conducted on equal-weighted portfolios⁵, which supports our findings.

Table 6 here

Table 7 reports the results of the spanning regression tests. Our starting point is the maximum Sharpe ratios in Panel A: They are highest for the FFL5F model (0.489), followed by the Carhart 4F (0.485) model. Next, the results across the spanning regression tests associated with all models are displayed in

⁴ The results are available in Table 2 in the online appendix.

⁵ The results are available in Table 3 in the online appendix.

Panels B to H. These reveal that *HML* lacks statistical strength and is susceptible to redundancy in asset pricing as argued by Fama and French (2018) using similar US data. This finding is attributable to an extremely low marginal contribution to the model's $Sh^2(f)$, i.e., $a^2/s^2(e)$ from *HML*, as well as a lack of statistical significance of the regression alpha. The next observation is a similarly weak level of statistical support for the *FFL* factor within the FFL2F framework. This *FFL* factor is formed from the returns difference between extreme decile portfolios following the sorting of the cross-section of stock returns by free float. The evidence alluding to the redundancy of the single-pass *FFL* factor is also built on the extremely low marginal contribution to the model's $Sh^2(f)$ and a lack of statistical significance of the regression alpha. However, in contrast to this single-pass *FFL* factor, there is markedly stronger statistical support for the size-adjusted *FFL** within multifactor configurations, all of which include the initial size factor. These are the FFL3F, FFL4F, and FFL5F models, respectively. Importantly, this evidence supports the relative importance of the size-adjusted *FFL** factor within multifactor asset pricing models.

Table 7 here

7. Conclusions

The motivations of insiders and controlling shareholders within a firm toward expropriating outside minority investors exert a profound influence on the degree to which the firm diversifies its ownership base (Berle and Means, 1932). We argue that within institutional environments characterized by high-quality protections afforded for external contracting, such as the US, controlling insiders are confronted with a trade-off. On the one hand, they can use their concentrated control to expropriate minority interests. On the other hand, there are potential penalties associated with elevated perceived risks of minority expropriation reflected in higher risk premiums and accompanying higher costs of equity. We argue that this conundrum reflects an overlooked state variable in asset pricing. We propose that the percentage free float is a useful metric in a single-country universe, and that this more theoretically driven asset pricing measure supersedes an evolving plethora of ratios and metrics based on various aspects of the firm's balance sheet and financial reporting.

We propose that a size-adjusted version of the free float derived from a double-sort process, based on an initial sort by size followed by a sort on free float, is a useful candidate to differentiate the cross-section of stock returns. Furthermore, this provides a means to create factors from the returns difference between extreme portfolios following stock sorting by size-adjusted free float. We argue that this yields a useful means for hedging effectively against the risks associated with the fundamental underlying likelihood of expropriation in a specific firm based on its ownership structure. Extensive

tests of variants of the augmented CAPM, including time-series, cross-sectional, and spanning regression tests, show that the CAPM augmented with both size and an FMP based on our new free-float metric yields the highest explanatory power for a cross-section of stock returns. The evidence supports our initial propositions after extensive testing and application to a large universe formed from the constituents of the US S&P 1500 benchmark index. Importantly, we argue that within institutional contexts supportive of arm's length contracting and external financing, there is increased importance in asset pricing on the size of listings and how dispersed as opposed to concentrated these are. This is important for minority outside investors who command a premium to compensate for potential welfare losses when buying into firms characterized by concentrated insider and controlling shareholder holdings.

References

- Albuquerque, R., Wang, N., 2008. Agency conflicts, investment, and asset pricing. *Journal of Finance*, 63(1), 1–40
- Anderson, R. C., Duru, A., Reeb, D. M., 2009. Founders, heirs, and corporate opacity in the United States. *Journal of Financial Economics*, 92(2), 205–222
- Anderson, R. C., Mansi, S. A., Reeb, D. M., 2003. Founding family ownership and the agency cost of debt. *Journal of Financial Economics*, 68(2), 263–285
- Ang, A., Hodrick, R., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. *Journal of Finance*, 51, 259–299
- Bae, K.-H., Baek, J.-S., Kang, J.-K., Liu, W.-L., 2012. Do controlling shareholders' expropriation incentives imply a link between corporate governance and firm value? Theory and evidence. *Journal of Financial Economics*, 105, 412–435
- Barillas, F., Shanken, J., 2017. Which alpha? *Review of Financial Studies*, 30, 1316–1338
- Barillas, F., Shanken, J., 2018. Comparing asset pricing models. *Journal of Finance*, 73, 715–754
- Bebchuck, L., Cohen, A., Ferrell, A., 2009. What matters in corporate governance. *Review of Financial Studies*, 22, 783–827
- Berle, A., Means, G., 1932. *The Modern Corporation and Private Property*. New York, NY: Macmillan
- Black, F., Jensen, M. C., Scholes, M., 1972. The capital asset pricing model: Some empirical tests. In M. C. Jensen (ed.), *Studies in the Theory of Capital Markets*. New York, NY: Praeger
- Carhart, M. M., 1997. On persistence in mutual fund performance. *Journal of Finance*, 52, 57–82
- Charitou, A., Louca, C., Panayides, S., 2007. Cross-listing, bonding hypothesis and corporate governance. *Journal of Business Finance & Accounting*, 34(7), 1281–1306
- Cochrane, J. H., 2005. *Asset Pricing*. Princeton, NJ: Princeton University Press
- Cooper, I., Ma, L., Maio, P., Philip, D., 2021. Multifactor models and their consistency with the APT. *Review of Asset Pricing Studies*, 11, 402–444
- Cooper, I., Maio, P., 2019a. Asset growth, probability, and investment opportunities. *Management Science*, 65, 3988–4010
- Cooper, I., Maio, P., 2019b. New evidence on conditional factor models. *Journal of Financial and Quantitative Analysis*, 54(5), 1975–2016
- Demsetz, H., Lehn, K., 1985. The structure of corporate ownership: Causes and consequences. *Journal of Political Economy*, 93(6), 1155–1177
- Dyck, A., Zingales, L., 2004. Private benefits of control: An international comparison. *Journal of Finance*, 59(2), 537–600

- Fama, E. F., 1996. Multifactor portfolio efficiency and multifactor asset pricing. *Journal of Financial and Quantitative Analysis*, 31(4), 441–465
- Fama, E. F., French, K. R., 1992. The cross section of expected stock returns. *The Journal of Finance*, 47, 427–465
- Fama, E. F., French, K. R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33, 3–56
- Fama, E. F., French, K. R., 1996. Multifactor explanations of asset pricing anomalies. *Journal of Finance*, 51, 55–84
- Fama, E. F., French, K. R., 2008. Dissecting anomalies. *Journal of Finance*, 63, 1653–1678
- Fama, E. F., French, K. R., 2015a. Incremental variables and the investment opportunity set. *Journal of Financial Economics*, 117, 470–488
- Fama, E. F., French, K. R., 2015b. A five-factor asset pricing model. *Journal of Financial Economics*, 116, 1–22
- Fama, E. F., French, K. R., 2016. Dissecting anomalies with a five-factor model. *Review of Financial Studies*, 29, 69–103
- Fama, E. F., French, K. R., 2018. Choosing factors. *Journal of Financial Economics*, 128, 234–252
- Fama, E. F., MacBeth, J. D., 1973. Risk, return & equilibrium: Empirical tests. *Journal of Political Economy*, 81(3), 607–636
- Frazzini, A., Pedersen, L. H., 2014. Betting against beta. *Journal of Financial Economics*, 111, 1–25
- Giannetti, M., Koskinen, Y., 2010. Investor protection, equity returns, and financial globalization. *Journal of Financial and Quantitative Analysis*, 45(1), 135–168
- Giannetti, M., Simonov, A., 2006. Which investors fear expropriation? Evidence from investors' portfolio choices. *Journal of Finance*, 61, 1507–1547
- Gibbons, M. R., Ross, S. A., Shanken, J., 1989. A test of the efficiency of a given portfolio. *Econometrica*, 57, 1121–1152
- Gompers, P., Ishii, J., Metrick, A., 2003. Corporate governance and equity prices. *The Quarterly Journal of Economics*, 118(1), 107–155
- Hearn, B., Phylaktis, K., Piesse, J., 2017. Expropriation risk by block holders, institutional quality and expected stock returns. *Journal of Corporate Finance*, 45, 122–149
- Helwege, J., Pirinsky, C., Stulz, R. M., 2007. Why do firms become widely held? An analysis of the dynamics of corporate ownership. *Journal of Finance*, LXII(3), 995–1028
- Hou, K., Karolyi, G. A., Kho, B.-C., 2011. What factors drive global stock returns? *Review of Financial Studies*, 24(8), 2527–2574

- Hou, K., Mo, H., Xue, C., Zhang, L., 2019. Which factors? *Review of Finance*, 23, 1–35
- Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. *Review of Financial Studies*, 28, 650–705
- Hou, K., Xue, C., Zhang, L., 2020. Replicating anomalies. *The Review of Financial Studies*, 33, 2019–2133
- Ikenberry, D., Lakonishok, J., Vermaelen, T., 1995. Market under reaction to open market share repurchases. *Journal of Financial Economics*, 39, 181–208
- Ince, O., Porter, R. B., 2003. Individual equity return data from Thomson Datastream: Handle with care! *Journal of Financial Research*, 29, 463–479
- Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1), 65–91
- Kan, R., Zhang, C. 1999. Two-pass tests of asset pricing models with useless factors. *The Journal of Finance*, 54, 203–235
- Lewellen, J., Nagel, S., Shanken, J., 2010. A skeptical appraisal of asset-pricing tests. *Journal of Financial Economics*, 96, 175–194
- Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47, 13–37
- Liu, W., 2006. A liquidity-augmented capital asset pricing model. *Journal of Financial Economics*, 82, 631–671
- Liu, J., Stambaugh, R. F., Yuan, Y., 2019. Size and value effects in China. *Journal of Financial Economics*, 134, 48–69
- Loughran, T., Ritter, J., 1995. The new issues puzzle. *Journal of Finance*, 50, 23–51
- MacKinlay, A. C., 1995. Multifactor models do not explain deviations from the CAPM. *Journal of Financial Economics*, 38, 3–28
- Maio, P., Santa-Clara, P., 2012. Multifactor models and their consistency with the ICAPM. *Journal of Financial Economics*, 106, 586–613
- Maio, P., Santa-Clara, P., 2017. Short-term interest rates and stock market anomalies. *Journal of Financial and Quantitative Analysis*, 52, 927–961
- Merton, R. C., 1973. An intertemporal capital asset pricing model. *Econometrica*, 41, 867–887
- Newey, W. K., West, K. D., 1987. A simple positive-definite heteroskedasticity and autocorrelation-consistent covariance matrix. *Econometrica*, 55, 703–708
- Roll, R., 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance*, 39, 1127–1139

- Ross, S. A., 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13, 341–360.
- Sharpe, W. F., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19, 425–442
- Sloan, R., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review*, 71, 289–315

Table 1 Fama-MacBeth (two-step procedure) regressions of stock returns on beta, size, and valuation ratios

The table reports the average slope coefficients from month-by-month Fama–MacBeth regressions for the US S&P 1500 market universe in the sample period of 1985:01 to 2023:12. Individual stock returns are regressed cross-sectionally on stock characteristics as of the previous month. The columns correspond to different regression specifications, with nonempty rows indicating the included regressors. λ_0 is the constant term. The regressors include pre-ranking CAPM $\beta^{pre-rank}$, estimated using the previous 60 months (over 5 years) of monthly returns; the log of the month-end market cap, *MCAP*, the book-to-market value ratio, *B/M*; and momentum, *MOM*, i.e., the time-series average of the percentage cumulative return for each stock over the previous 12 months, omitting the most recent month, and is monthly, following Jegadeesh and Titman (1993). Also included is Liu’s (2006) liquidity measure, *LIU*, estimated over the previous one-year ranking period, and the percentage free float of the listed market capitalization, *Free Float*. The last row reports the average R^2 for each specification. Panel A reports for the aggregate US S&P 1500 market universe constituents, while Panel B reports for the constituents within the three smallest size portfolios following annual sorting into 10 decile size-based portfolios, with Panel C reporting on the constituents of the three largest size-sorted portfolios. The t-statistics based on Newey and West (1987) standard errors with four lags are reported in parentheses.

	λ_0	$\beta^{pre-rank}$	MCAP	B/M	MOM	LIU	Free Float	R^2
Panel A: Overall								
(1)	0.013 [5.58]	0.001 [1.14]	-- --	-- --	-- --	-- --	-- --	0.005
(2)	0.054 [6.29]	0.001 [1.15]	-0.002 [-5.29]	-0.007 [-4.83]	-- --	-- --	-- --	0.028
(3)	0.042 [5.61]	0.001 [0.99]	-0.001 [-4.80]	-0.005 [-3.47]	0.018 [4.74]	-- --	-- --	0.062
(4)	0.019 [5.98]	0.001 [0.66]	-- --	-- --	-- --	-0.003 [-4.23]	-- --	0.015
(5)	0.018 [5.45]	0.002 [1.43]	-- --	-- --	-- --	-- --	-0.007 [-3.29]	0.009
(6)	0.043 [4.62]	0.002 [1.34]	-0.001 [-3.43]	-- --	-- --	-- --	-0.006 [-3.90]	0.021
(7)	0.056 [6.48]	0.001 [1.35]	-0.002 [-5.13]	-0.007 [-4.88]	-- --	-- --	-0.005 [-3.18]	0.033
(8)	0.044 [5.73]	0.001 [1.17]	-0.001 [-4.75]	-0.005 [-3.67]	0.0175 [4.65]	-- --	-0.002 [-1.48]	0.068
Panel B: Smallest (D1, D2, D3)								
(1a)	0.022 [7.54]	0.001 [0.54]	-- --	-- --	-- --	-- --	-- --	0.006
(2a)	-0.005 [-0.26]	0.001 [1.09]	0.001 [1.55]	-0.008 [-4.97]	-- --	-- --	-- --	0.030
(3a)	0.002 [0.14]	0.001 [0.97]	0.001 [0.97]	-0.006 [-4.27]	0.013 [4.06]	-- --	-- --	0.061
(4a)	0.027 [6.94]	0.001 [0.41]	-- --	-- --	-- --	-0.004 [-2.95]	-- --	0.018
(5a)	0.022 [4.54]	0.001 [1.23]	-- --	-- --	-- --	-- --	-0.002 [-1.49]	0.015
(6a)	-0.030 [-1.30]	0.001 [1.17]	0.003 [2.30]	-- --	-- --	-- --	-0.004 [-1.82]	0.026
(7a)	-0.003 [-0.14]	0.001 [1.18]	0.001 [1.44]	-0.008 [-4.50]	-- --	-- --	-0.003 [-1.69]	0.046
(8a)	0.002 [0.10]	0.001 [0.90]	0.001 [0.81]	-0.006 [-3.76]	0.014 [4.32]	-- --	-0.002 [-1.42]	0.080
Panel C: Biggest (D8, D9, D10)								
(1b)	0.008 [3.57]	0.001 [0.94]	-- --	-- --	-- --	-- --	-- --	0.014
(2b)	-0.027 [-3.01]	0.001 [0.96]	0.001 [4.81]	-0.008 [-3.78]	-- --	-- --	-- --	0.045
(3b)	-0.020 [-2.68]	0.001 [0.84]	0.001 [4.42]	-0.005 [-3.11]	0.022 [4.28]	-- --	-- --	0.101
(4b)	0.008 [2.33]	0.001 [0.83]	-- --	-- --	-- --	-0.003 [-0.36]	-- --	0.027
(5b)	0.009 [2.47]	0.001 [1.12]	-- --	-- --	-- --	-- --	-0.001 [-1.29]	0.022
(6b)	-0.041 [-3.45]	0.001 [1.24]	0.002 [4.87]	-- --	-- --	-- --	-0.002 [-1.92]	0.032
(7b)	-0.032 [-2.94]	0.001 [1.23]	0.001 [4.53]	-0.007 [-3.56]	-- --	-- --	-0.001 [-1.18]	0.056
(8b)	-0.025 [-2.69]	0.001 [0.89]	0.001 [3.98]	-0.005 [-3.08]	0.021 [3.89]	-- --	-0.001 [-1.51]	0.112

Table 2 Factor mimicking portfolio aggregate descriptive statistics

This table reports the descriptive statistics, correlations, and autocorrelations (at 1, 6, and 12 lags) for the sample, which is the US S&P 1500 market universe for the period 1985:01 to 2023:12 for returns-based valuation factors, including the market excess returns, the Fama and French (1993) size (*SMB*) and book-to-market value (*HML*), and momentum factor, namely, up minus down (*UMD*). Also included are Liu's (2006) liquidity measure (*ILLIQ*) and the new factor formed on free float (*FFL*), which is in two forms: the first (*FFL*) being the difference portfolio between low and high extreme decile portfolios and the second (*FFL**) being from the second stage sort of a 3×3 double-sort procedure first sorting on size, or market capitalization, into three tercile portfolios and then each subsequently sorted into a further three tercile portfolios based on free float. The resulting second stage size-adjusted free float factor is formed from average returns on three lowest free float tercile portfolios minus the average returns from three highest free float tercile portfolios. The market universe is the aggregate US S&P 1500 and, in addition to all factors, is value weighted. The *SMB* and *HML* factors are formed by a 3×3 double-sort procedure with an initial sort into five-tercile portfolios based on market capitalization or size followed by a secondary sort based on the book-to-market value for *HML*. The procedure for forming the *ILLIQ* and *FFL* factors is different, as it relies on sorting into 10-decile portfolios with the hedging factor resulting from the difference between the extreme portfolios. Descriptive statistics reported in Panel A include the monthly average or mean returns in addition to the t-statistic indicating the significance from zero, standard deviation, skewness and kurtosis in distribution, Jarque-Bera statistics for non-normality, and the sample period in number of months. Correlations between factors are reported in Panel B, while autocorrelations are reported in Panel C. †, *, and ** indicate significance at the 10%, 5%, and 1% levels, respectively.

	R_m-R_f	SMB	HML	UMD	ILLIQ	FFL	FFL*
Panel A: Descriptive statistics							
Mean (%)	0.87	1.30	-0.15	0.84	-1.59	0.20	0.05
t-Statistic	2.29	8.04	-0.86	3.80	-4.77	1.62	0.57
Standard deviation (%)	7.65	3.28	3.74	4.49	6.74	4.10	1.94
Skewness	-0.09	0.23	0.07	-0.52	-0.24	0.45	-0.26
Kurtosis	7.24	3.39	5.65	7.06	4.52	9.04	5.23
Jarque-Bera statistic [<i>p</i> -value]	352.33 [0.00]	7.26 [0.02]	137.06 [0.00]	342.88 [0.00]	49.48 [0.00]	726.75 [0.00]	102.49 [0.00]
Number of months	468	468	468	468	468	468	468
Panel B: Pearson correlations							
Market	1.00						
SMB	-0.14**	1.00					
HML	-0.21**	0.21**	1.00				
UMD	-0.06†	-0.08*	-0.36**	1.00			
ILLIQ	0.06†	-0.38**	0.25**	-0.10*	1.00		
FFL	0.13**	0.08*	-0.22**	0.11*	-0.20**	1.00	
FFL*	0.23**	0.09*	-0.08*	0.01	-0.04	0.39**	1.00
Panel C: Autocorrelations							
1-Lag	-0.05	0.02	0.08	0.15*	-0.05	0.04	0.12*
6-Lags	0.04	-0.06	0.02*	0.12*	0.16*	0.06	0.07*
12-Lags	-0.16**	-0.02	-0.03**	0.04**	-0.07*	-0.15*	0.06*

Table 3 Factor mimicking portfolio characteristic summary statistics

This table reports the stock counts, summary descriptive statistics, and categories of block ownership for each of the 10 value-weighted free float sorted decile portfolios (D1–D10). The sample is for the US S&P 1500 market universe for the period 1985:01 to 2023:12. Panel A reports the stock counts per index constituency for the US, namely, the S&P 500 [large cap], S&P 400 [mid-cap], S&P 600 [small cap], and the S&P 1500 overall. Panel B reports the average equal and value-weighted returns followed by the average proportions of *Free Float* (free float, %), *MOM* (momentum change over 12 months, %), *B/M* (book-to-market ratio), the trading statistics of *MCAP* (market capitalization, US\$ billions), *Volume* (daily traded volume, US\$ millions), and *Price* (stock price – an indicator of risk, US\$), and the anomaly variables of *Beta* (β), *NSI* (change in net stock issues from year 2 to year 1), *Volatility* (daily price volatility, %), *Asset Growth* (5-year asset growth, year 5 – year 1), *Accrual* (two-year change in accruals – (year 2 – year 1) scaled by book equity), and *OP* (operating profit again scaled by book equity). In the first column of Panel B, a t-difference in the means' statistical significance confidence level is provided for the mean values in decile portfolio D1 in relation to the differences between these values and those for D10. †, *, and ** indicate significance at the 10%, 5%, and 1% levels, respectively. Panel C reports average ownership per block owner category reported by Refinitiv Datastream, namely, cross-shareholder networks, employee/family, foreign, state, investment companies, and other and pension funds.

	D1 (Low)	D2	D3	D4	D5	D6	D7	D8	D9	D10 (High)
Panel A: Stock count (#)										
US S&P 500 [Large cap]	22.42	24.71	28.40	36.76	34.26	36.59	38.21	49.95	56.64	72.27
US S&P 400 [Mid-cap]	27.38	26.07	27.45	27.23	26.66	28.15	27.81	30.84	22.98	28.57
US S&P 600 [Small cap]	63.60	59.59	55.13	47.94	50.17	46.94	43.03	31.30	31.75	11.44
Total	113.41	110.37	110.97	111.93	111.10	111.69	109.04	112.08	111.37	112.29
Panel B: Summary statistics										
Returns – value weight (%)	1.36†	1.51	0.71	1.13	1.07	1.56	1.22	1.16	1.12	1.20
Returns – equal weight (%)	1.73†	1.54	1.52	1.51	1.47	1.52	1.49	1.45	1.39	1.31
Free Float	58.48**	75.62	80.35	84.53	87.14	89.32	90.97	92.58	94.03	94.66
MOM	17.84**	24.02	12.45	13.65	12.59	23.35	14.88	13.33	10.32	11.25
B/M	0.38**	0.42	0.55	0.43	0.41	0.44	0.48	0.44	0.51	0.40
MCAP	98.36**	215.28	280.72	211.33	290.20	143.81	536.87	191.74	944.51	298.01
Volume	79.68**	66.09	122.99	92.37	69.21	97.13	136.44	128.94	96.80	180.63
Price	38.70**	85.25	56.35	36.89	75.10	83.36	110.39	58.31	254.24	87.26
LIU	581.88	2,713.41	963.00	32.27	776.80	1,186.10	312.35	209.92	6,494.78	731.69
β	1.17**	1.12	1.07	1.08	1.09	1.11	1.08	1.09	0.99	1.01
NSI	0.09†	0.09	0.08	0.08	0.10	0.09	0.10	0.10	0.07	0.07
Volatility	0.02**	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.01
Asset Growth	13.39**	14.73	11.93	12.42	11.69	13.15	12.70	13.54	11.07	9.71
Accrual	0.08**	-0.12	0.03	0.06	0.01	0.24	-0.01	-0.04	0.02	-0.01
OP	0.26**	0.22	0.20	0.23	0.28	0.03	0.11	0.24	0.31	-0.02

Panel C: Block ownership

Cross-shareholder networks (%)	2.43**	0.66	0.69	0.81	1.02	1.23	1.02	1.15	0.82	0.73
Employee/Family (%)	2.95**	3.41	3.78	1.97	1.76	2.46	2.20	1.28	6.62	1.26
Foreign (%)	0.91**	0.84	0.99	1.03	0.71	0.82	0.93	0.69	0.84	0.71
State (%)	0.12**	0.05	0.06	0.04	0.02	0.02	0.01	0.07	0.08	0.07
Investment companies (%)	8.52	9.22	11.33	9.20	9.42	7.74	6.59	8.59	6.65	8.11
Other (%)	0.09**	0.30	0.12	0.43	0.10	0.18	0.24	0.20	0.31	0.27
Pension funds (%)	0.05**	0.27	0.04	0.04	0.32	0.09	0.09	0.04	0.40	0.12

Table 4 Regression results for 10 free float decile portfolios

This table reports the results for the time-series regression beta coefficients for valuation factors with t-statistics and explanatory power (R^2) for the four asset pricing models that do not include and free float factors. The sample is for the US S&P 1500 market universe for the period 1985:01 to 2023:12. The dependent variables are each of the value-weighted 10-decile free float–sorted portfolio returns in addition to a final high-minus-low difference portfolio formed from the difference of D10 and D1 free float (*FFL*) portfolio returns. D1 is the lowest free float (most concentrated block ownership), and D10 the highest free float (most dispersed ownership). Panel A reports the capital asset pricing model (CAPM) comprising only the market factor denominated in excess returns. Panel B reports the Fama and French (1993) three factor model, FF3F, comprising the CAPM plus *SMB* (small minus big) and *HML* (high minus low) factors, related to the size and the book-to-market value. Panel C reports the Carhart (1997) four factor model, Carhart 4F, comprising the FF3F factors plus an additional momentum factor, *UMD* (up minus down) factor. Panel D outlines Liu’s (2006) two-factor model, ILLIQ2F, comprising the CAPM plus by an illiquidity factor (*ILLIQ*) formed from the Liu (2006) liquidity metric. The market universe is the aggregate US S&P 1500, with all factors being value weighted. The *SMB*, *HML*, and *UMD* factors are all formed by a 3×3 double-sort procedure with an initial sort into five-tercile portfolios based on market capitalization or size followed by a secondary sort based on each of the additional factors, namely, the book-to-market value for *HML* or momentum for *UMD*. The procedure for forming the *ILLIQ* and *FFL* factors is different, as it relies on sorting into 10-decile portfolios with hedging factors resulting from the difference between the extreme portfolios. Rebalancing in all cases is annual in December of each year. The one-month US deposit yield is used as the risk-free rate. Numbers in square brackets are t-statistics. †, *, and ** indicate significance at the 10%, 5%, and 1% levels, respectively. Standard errors are Huber-White heteroscedasticity-robust.

	D1 (Low)	D2	D3	D4	D5	D6	D7	D8	D9	D10 (High)	High minus Low
Panel A: CAPM											
Alpha (%)	0.007 [3.41]	0.008 [2.60]	0.001 [0.03]	0.005 [2.18]	0.004 [1.74]	0.008 [2.55]	0.006 [2.74]	0.005 [2.46]	0.005 [2.06]	0.006 [3.64]	-0.004 [-2.18]
Beta: Market (excess return)	0.391 [8.95]	0.525 [9.01]	0.490 [8.30]	0.429 [8.80]	0.457 [8.16]	0.581 [8.26]	0.436 [9.41]	0.411 [8.93]	0.412 [8.64]	0.320 [8.23]	-0.070 [-2.96]
R^2	0.297	0.276	0.278	0.334	0.362	0.299	0.379	0.333	0.297	0.307	0.018
Panel B: FF3F											
Alpha (%)	0.003 [2.15]	0.001 [1.21]	-0.007 [-2.43]	-0.001 [-1.24]	0.001 [1.22]	0.003 [0.98]	0.001 [0.41]	0.001 [0.39]	0.003 [2.10]	0.004 [2.05]	-0.004 [-2.20]
Beta: Market (excess return)	0.405 [9.68]	0.553 [10.28]	0.566 [10.43]	0.466 [10.34]	0.498 [9.38]	0.622 [9.06]	0.467 [10.87]	0.441 [10.24]	0.456 [10.24]	0.348 [9.57]	-0.048 [-2.09]
Beta: SMB	0.353 [4.18]	0.523 [3.96]	0.530 [5.85]	0.400 [4.82]	0.268 [3.14]	0.357 [3.06]	0.361 [5.12]	0.328 [4.38]	0.166 [2.09]	0.173 [2.63]	0.009 [0.11]
Beta: HML	-0.070 [-0.88]	-0.043 [-0.22]	0.424 [4.09]	0.127 [1.55]	0.240 [3.19]	0.187 [1.13]	0.093 [1.34]	0.101 [1.31]	0.331 [4.20]	0.176 [2.84]	0.220 [3.15]
R^2	0.337	0.321	0.397	0.397	0.413	0.330	0.432	0.378	0.352	0.348	0.056

Panel C: Carhart 4F											
Alpha (%)	0.003	-0.001	-0.007	0.001	-0.001	-0.001	0.001	0.001	0.003	0.004	-0.003
	[2.14]	[-0.19]	[-2.49]	[0.10]	[-0.22]	[-0.30]	[0.29]	[1.43]	[2.16]	[2.18]	[-1.91]
Beta: Market (excess return)	0.406	0.568	0.569	0.458	0.509	0.666	0.470	0.440	0.454	0.344	-0.056
	[9.62]	[10.5]	[10.55]	[10.11]	[9.70]	[9.93]	[11.14]	[10.18]	[10.38]	[9.59]	[-2.30]
Beta: SMB	0.353	0.528	0.532	0.396	0.272	0.374	0.362	0.327	0.165	0.172	0.039
	[4.19]	[4.01]	[5.86]	[4.81]	[3.28]	[3.27]	[5.10]	[4.36]	[2.08]	[2.60]	[0.42]
Beta: HML	-0.068	0.040	0.442	0.080	0.303	0.427	0.108	0.093	0.321	0.154	0.197
	[-0.80]	[0.23]	[3.97]	[0.88]	[3.45]	[2.26]	[1.35]	[1.09]	[3.54]	[2.36]	[2.93]
Beta: UMD	0.005	0.173	0.036	-0.099	0.130	0.500	0.032	-0.015	-0.020	-0.045	-0.057
	[0.08]	[1.62]	[0.46]	[-1.23]	[1.83]	[3.01]	[0.50]	[-0.23]	[-0.27]	[-0.90]	[-0.96]
R ²	0.337	0.329	0.397	0.402	0.421	0.391	0.433	0.378	0.352	0.350	0.059
Panel D: ILLIQ2F											
Alpha (%)	0.004	0.003	-0.002	0.001	0.002	0.003	0.003	0.002	0.004	0.005	-0.002
	[1.78]	[2.22]	[-1.74]	[0.63]	[0.78]	[1.14]	[1.23]	[1.80]	[1.62]	[2.49]	[-1.68]
Beta: Market (excess return)	0.405	0.542	0.499	0.442	0.465	0.598	0.448	0.425	0.415	0.327	-0.077
	[8.43]	[8.43]	[8.05]	[8.23]	[7.99]	[7.84]	[8.90]	[8.47]	[8.43]	[7.81]	[-3.24]
Beta: ILLIQ	-0.239	-0.289	-0.150	-0.225	-0.143	-0.297	-0.206	-0.231	-0.048	-0.118	0.121
	[-6.34]	[-3.47]	[-3.03]	[-5.94]	[-3.88]	[-5.21]	[-5.54]	[-6.20]	[-1.06]	[-3.77]	[3.35]
R ²	0.388	0.344	0.299	0.410	0.391	0.364	0.449	0.419	0.300	0.341	0.063

Table 5 Differentiation between models applied to anomalies in cross-section of stock returns

This table presents the Gibbons, Ros, and Shanken (GRS; 1989) statistics for the US S&P 1500 market universe for the period 1985:01 to 2023:12 for each of the eight asset pricing models as applied to 10-decile portfolios formed from sorting of stocks by 14 anomaly variables. These are *Free Float*, pre-ranking Beta ($\beta_{pre-rank}$), market capitalization (*MCAP*), accounting book-to-market value (*B/M*), *Volatility*, *Asset Growth* over preceding five years, *Accrual* scaled by book equity, Operating profit (*OP*), Dividend yield (*DY*), Price to cash flow (*P/CF*), Net stock issues (*NSI*), Price-to-earnings ratio (*P/E*), Growth in capital expenditure (*CAPX Growth*), and *Sales Growth* over preceding one year. The eight asset pricing models are the CAPM, FF3F (including the additional SMB and HML), Carhart 4F (including the additional momentum factor, UMD), the two-factor liquidity augmented CAPM by the Liu (2006) liquidity hedging portfolio, namely, ILLIQ2F, and the similar two-factor augmented CAPM with free float (formed from decile sorting of 10 portfolios followed by returns difference between low-minus-high), namely, FFL2F. These are followed by FFL3F, which is CAPM plus *SMB* and *FFL** both formed from 3×3 double-sort procedure, then FFL4F, which is CAPM plus *SMB*, *HML*, and *FFL**, all formed from a 3×3 double-sort procedure with *SMB* being the average of the underlying two size factors formed from each of the two component double sorts, and finally FFL5F, comprising CAPM plus *SMB*, *HML*, *UMD*, and *FFL** with these formed from three individual 3×3 double-sort procedures and size factor being the average of each constituent size factor from each double sort. The GRS statistic tests whether all intercepts in a set of test portfolios (assets) regressions are zero, with the associated *p*-value in square brackets, MAA is the average absolute intercept for a set of regressions, R_C^2 is the cross-sectional constrained R^2 and the numbers in parentheses represent the respective empirical *p*-values to test the null that the explanatory ratio is zero (obtained from a bootstrap simulation), and finally, R^2 is the average adjusted- R^2 . The regression alpha (constant) and associated *t*-statistic are reported for the high-minus-low hedging or difference portfolios for each of the 14 anomaly variables and are also reported in the farthest two right-hand columns for each of the eight asset pricing models. Standard errors are Huber-White heteroscedasticity-robust.

	Ten decile anomaly portfolios				Hedging anomaly portfolio	
	Free Float				Alpha (%)	[t-statistic]
	R^2	R_C^2 [p-value]	MAA, %	GRS [p-value]		
CAPM	0.316	0.086 [0.10]	0.005	26.64 [0.00]	-0.003	-2.13
FF3F	0.371	-0.619 [0.85]	0.002	26.61 [0.00]	-0.001	-0.67
Carhart 4F	0.379	-0.692 [0.86]	0.002	31.27 [0.00]	-0.001	-0.46
ILLIQ2F	0.403	0.284 [0.02]	0.002	19.40 [0.00]	-0.002	-1.12
FFL2F	0.360	0.159 [0.04]	0.005	24.73 [0.00]	-0.002	-28.05
FFL3F	0.380	-0.547 [0.83]	0.002	29.70 [0.00]	-0.003	-1.77
FFL4F	0.395	-0.193 [0.51]	0.002	27.05 [0.00]	-0.002	-1.37
FFL5F	0.403	-0.221 [0.52]	0.002	28.44 [0.00]	-0.002	-1.15
	$\beta_{pre-rank}$					
	R^2	R_C^2 [p-value]	MAA, %	GRS [p-value]	Alpha (%)	[t-statistic]
CAPM	0.310	-0.003 [0.45]	0.004	24.12 [0.00]	-0.001	-0.50
FF3F	0.374	-1.021 [0.95]	0.002	18.07 [0.00]	-0.009	-2.53
Carhart 4F	0.375	-0.796 [0.88]	0.001	19.72 [0.00]	-0.008	-2.32
ILLIQ2F	0.345	0.110 [0.17]	0.003	22.13 [0.00]	-0.006	-1.98

FFL2F	0.318	0.012 [0.38]	0.004	12.31 [0.00]	-0.002	-0.60
FFL3F	0.351	-0.788 [0.90]	0.001	17.52 [0.00]	-0.008	-2.42
FFL4F	0.382	-0.603 [0.83]	0.001	18.45 [0.00]	-0.007	-2.23
FFL5F	0.382	-0.371 [0.64]	0.001	17.67 [0.00]	-0.007	-2.01

MCAP

	R ²	R _C ² [p-value]	MAA, %	GRS [p-value]	Alpha (%)	[t-statistic]
CAPM	0.425	-0.079 [0.81]	0.012	123.68 [0.00]	-0.035	-10.82
FF3F	0.649	0.765 [0.00]	0.002	51.77 [0.00]	-0.015	-6.33
Carhart 4F	0.652	0.765 [0.00]	0.002	58.52 [0.00]	-0.015	-6.14
ILLIQ2F	0.497	0.113 [0.14]	0.009	118.33 [0.00]	-0.031	-9.69
FFL2F	0.432	-0.064 [0.70]	0.012	45.69 [0.00]	-0.034	-10.82
FFL3F	0.636	0.768 [0.00]	0.002	59.16 [0.00]	-0.014	-5.86
FFL4F	0.651	0.749 [0.00]	0.002	65.60 [0.00]	-0.015	-6.49
FFL5F	0.653	0.747 [0.00]	0.002	66.61 [0.00]	-0.015	-6.30

B/M

	R ²	R _C ² [p-value]	MAA, %	GRS [p-value]	Alpha (%)	[t-statistic]
CAPM	0.365	0.191 [0.10]	0.004	19.83 [0.00]	0.005	1.24
FF3F	0.468	0.297 [0.05]	0.002	12.34 [0.00]	0.004	1.10
Carhart 4F	0.471	0.472 [0.01]	0.002	14.15 [0.00]	0.002	0.54
ILLIQ2F	0.395	0.040 [0.28]	0.002	15.03 [0.00]	0.007	1.82
FFL2F	0.370	0.160 [0.04]	0.004	17.71 [0.00]	0.005	1.36
FFL3F	0.407	0.473 [0.06]	0.002	11.78 [0.00]	-0.001	-0.16
FFL4F	0.470	0.207 [0.09]	0.002	15.80 [0.00]	0.004	1.15
FFL5F	0.474	0.411 [0.02]	0.002	15.38 [0.00]	0.002	0.59

Volatility

	R ²	R _C ² [p-value]	MAA, %	GRS [p-value]	Alpha (%)	[t-statistic]
CAPM	0.340	0.615 [0.00]	0.004	15.05 [0.00]	0.003	1.50
FF3F	0.381	0.218 [0.11]	0.002	13.25 [0.00]	-0.008	-1.53
Carhart 4F	0.383	0.038 [0.26]	0.002	14.72 [0.00]	-0.009	-1.67
ILLIQ2F	0.389	0.410 [0.01]	0.003	14.99 [0.00]	-0.006	-1.32
FFL2F	0.346	0.655 [0.00]	0.004	14.00 [0.00]	0.006	1.38
FFL3F	0.373	0.754 [0.02]	0.002	11.40 [0.00]	-0.002	-1.04
FFL4F	0.386	0.349 [0.04]	0.002	13.39 [0.00]	-0.007	-1.24
FFL5F	0.388	0.199 [0.12]	0.002	13.06 [0.00]	-0.007	-1.37

Asset Growth						
	R^2	R_c^2 [p -value]	MAA, %	GRS [p -value]	Alpha (%)	[t -statistic]
CAPM	0.327	0.158 [0.03]	0.005	24.47 [0.00]	-0.005	-1.51
FF3F	0.392	0.032 [0.24]	0.002	20.17 [0.00]	-0.007	-1.87
Carhart 4F	0.399	0.267 [0.06]	0.002	19.65 [0.00]	-0.003	-0.97
ILLIQ2F	0.388	-0.238 [0.82]	0.003	30.58 [0.00]	-0.010	-2.77
FFL2F	0.342	0.131 [0.06]	0.005	17.45 [0.00]	-0.006	-1.94
FFL3F	0.365	0.404 [0.01]	0.001	16.31 [0.00]	-0.002	-0.74
FFL4F	0.399	0.117 [0.15]	0.002	23.65 [0.00]	-0.005	-1.47
FFL5F	0.406	0.272 [0.05]	0.002	21.51 [0.00]	-0.003	-0.51
Accrual						
	R^2	R_c^2 [p -value]	MAA, %	GRS [p -value]	Alpha (%)	[t -statistic]
CAPM	0.342	-0.136 [0.91]	0.004	18.59 [0.00]	-0.007	-2.12
FF3F	0.398	0.153 [0.12]	0.001	7.90 [0.00]	-0.008	-2.64
Carhart 4F	0.402	0.075 [0.18]	0.001	11.04 [0.00]	-0.008	-2.37
ILLIQ2F	0.398	-1.142 [1.00]	0.002	18.15 [0.00]	-0.011	-3.69
FFL2F	0.352	-0.142 [0.89]	0.004	17.55 [0.00]	-0.007	-2.19
FFL3F	0.377	0.328 [0.03]	0.001	7.19 [0.00]	-0.006	-1.29
FFL4F	0.402	0.142 [0.12]	0.001	9.57 [0.00]	-0.008	-2.49
FFL5F	0.405	0.043 [0.21]	0.001	11.87 [0.00]	-0.007	-2.22
OP						
	R^2	R_c^2 [p -value]	MAA, %	GRS [p -value]	Alpha (%)	[t -statistic]
CAPM	0.347	-0.260 [0.98]	0.004	15.48 [0.00]	0.002	0.41
FF3F	0.400	-1.926 [1.00]	0.001	14.33 [0.00]	0.003	0.88
Carhart 4F	0.401	-2.205 [1.00]	0.001	16.15 [0.00]	0.002	0.67
ILLIQ2F	0.399	-3.029 [1.00]	0.002	11.48 [0.00]	0.005	1.54
FFL2F	0.354	-0.359 [0.99]	0.004	15.88 [0.00]	0.002	0.52
FFL3F	0.382	-2.585 [1.00]	0.001	20.83 [0.00]	0.003	0.79
FFL4F	0.403	-1.575 [0.99]	0.001	16.58 [0.00]	0.002	0.63
FFL5F	0.404	-1.891 [0.99]	0.001	16.78 [0.00]	0.001	0.41
DY						
	R^2	R_c^2 [p -value]	MAA, %	GRS [p -value]	Alpha (%)	[t -statistic]
CAPM	0.338	0.619 [0.00]	0.004	12.71 [0.00]	-0.002	-0.56
FF3F	0.411	-0.077 [0.38]	0.002	6.76 [0.00]	0.002	0.51

Carhart 4F	0.414	-0.917 [0.91]	0.002	11.45 [0.00]	0.004	0.87
ILLIQ2F	0.375	0.276 [0.03]	0.003	11.21 [0.00]	-0.001	-0.25
FFL2F	0.347	0.638 [0.00]	0.004	10.94 [0.00]	-0.002	-0.55
FFL3F	0.376	0.217 [0.09]	0.001	6.98 [0.00]	0.001	0.33
FFL4F	0.414	0.103 [0.20]	0.002	6.23 [0.00]	0.001	0.22
FFL5F	0.417	-0.687 [0.83]	0.002	9.90 [0.00]	0.003	0.67

P/CF

	R ²	R _C ² [p-value]	MAA, %	GRS [p-value]	Alpha (%)	[t-statistic]
CAPM	0.336	-1.497 [1.00]	0.005	19.05 [0.00]	0.005	1.50
FF3F	0.412	-0.641 [0.93]	0.002	9.49 [0.00]	0.002	0.76
Carhart 4F	0.413	-1.407 [0.98]	0.002	13.19 [0.00]	0.003	0.78
ILLIQ2F	0.381	-3.156 [1.00]	0.004	19.99 [0.00]	0.003	0.89
FFL2F	0.348	-1.426 [1.00]	0.005	15.34 [0.00]	0.005	1.40
FFL3F	0.375	-1.651 [0.99]	0.002	12.90 [0.00]	0.006	1.65
FFL4F	0.417	-0.640 [0.86]	0.002	9.22 [0.00]	0.001	0.68
FFL5F	0.418	-1.119 [0.95]	0.002	12.10 [0.00]	0.003	0.91

NSI

	R ²	R _C ² [p-value]	MAA, %	GRS [p-value]	Alpha (%)	[t-statistic]
CAPM	0.349	-0.383 [0.99]	0.004	22.68 [0.00]	-0.007	-3.48
FF3F	0.405	-0.149 [0.46]	0.001	8.94 [0.00]	-0.004	-2.01
Carhart 4F	0.410	0.037 [0.21]	0.001	9.00 [0.00]	-0.004	-1.97
ILLIQ2F	0.404	-1.289 [1.00]	0.002	25.00 [0.00]	-0.009	-4.21
FFL2F	0.360	-0.401 [0.99]	0.004	16.53 [0.00]	-0.008	-3.82
FFL3F	0.388	0.129 [0.14]	0.001	6.93 [0.00]	-0.003	-1.16
FFL4F	0.410	-0.180 [0.50]	0.001	10.05 [0.00]	-0.004	-1.77
FFL5F	0.415	0.019 [0.23]	0.001	9.04 [0.00]	-0.004	-1.73

P/E

	R ²	R _C ² [p-value]	MAA, %	GRS [p-value]	Alpha (%)	[t-statistic]
CAPM	0.323	-0.079 [0.81]	0.004	22.62 [0.00]	-0.003	-0.96
FF3F	0.396	-0.158 [0.47]	0.001	12.96 [0.00]	-0.001	-0.34
Carhart 4F	0.397	-0.493 [0.75]	0.002	17.50 [0.00]	-0.001	-0.28
ILLIQ2F	0.359	-0.166 [0.72]	0.003	19.10 [0.00]	-0.006	-1.96
FFL2F	0.328	-0.101 [0.81]	0.004	17.76 [0.00]	-0.003	-1.13
FFL3F	0.368	-0.710 [0.89]	0.001	17.58 [0.00]	0.002	0.66
FFL4F	0.403	-0.012 [0.40]	0.001	12.52 [0.00]	-0.001	-0.02

FFL5F	0.404	-0.382 [0.66]	0.002	16.99 [0.00]	0.001	0.04
-------	-------	---------------	-------	--------------	-------	------

CAPX Growth

	R ²	R _C ² [p-value]	MAA, %	GRS [p-value]	Alpha (%)	[t-statistic]
CAPM	0.332	-0.096 [0.85]	0.004	12.48 [0.00]	-0.008	-2.74
FF3F	0.382	-1.027 [0.96]	0.001	11.73 [0.00]	-0.005	-1.63
Carhart 4F	0.386	-0.418 [0.70]	0.001	11.15 [0.00]	-0.003	-0.94
ILLIQ2F	0.399	-1.269 [1.00]	0.002	14.11 [0.00]	-0.009	-3.00
FFL2F	0.340	-0.123 [0.86]	0.004	10.05 [0.00]	-0.008	-2.78
FFL3F	0.369	-0.469 [0.79]	0.001	10.31 [0.00]	-0.004	-1.26
FFL4F	0.388	-0.773 [0.91]	0.001	12.28 [0.00]	-0.005	-1.50
FFL5F	0.392	-0.020 [0.50]	0.001	10.02 [0.00]	-0.002	-0.79

Sales Growth

	R ²	R _C ² [p-value]	MAA, %	GRS [p-value]	Alpha (%)	[t-statistic]
CAPM	0.318	0.190 [0.01]	0.005	13.11 [0.00]	-0.003	-1.16
FF3F	0.388	0.347 [0.03]	0.001	7.24 [0.00]	-0.004	-1.36
Carhart 4F	0.389	0.311 [0.04]	0.001	9.18 [0.00]	-0.003	-0.94
ILLIQ2F	0.372	0.034 [0.29]	0.002	11.79 [0.00]	-0.007	-2.43
FFL2F	0.326	0.176 [0.02]	0.005	8.48 [0.00]	-0.004	-1.43
FFL3F	0.356	0.858 [0.00]	0.001	1.57 [0.00]	-0.001	-0.54
FFL4F	0.393	0.492 [0.00]	0.001	7.26 [0.00]	-0.003	-1.04
FFL5F	0.394	0.421 [0.02]	0.001	8.42 [0.00]	-0.002	-0.60

Table 6 Joint time-series tests: cross-sectional analysis

This table presents the results for the joint time-series tests of the conditional factor models for the US S&P 1500 market universe for the period 1985:01 to 2023:12. The test portfolios are the 14 sets of 10-decile portfolios for each of the 14 anomalies. The eight asset pricing models are the CAPM, FF3F (including the additional *SMB* and *HML*), Carhart 4F (including the additional momentum factor, *UMD*), the two-factor liquidity augmented CAPM by the Liu (2006) liquidity hedging portfolio, namely, ILLIQ2F, and the similar two-factor augmented CAPM with free float (formed from decile sorting of 10 portfolios followed by returns difference between low-minus-high), namely, FFL2F. These are followed by FFL3F, which is CAPM plus *SMB* and *FFL** (both formed from a 3×3 double-sort procedure), then FFL4F, which is CAPM plus *SMB*, *HML*, and *FFL**, all formed from 3×3 double-sort procedure with *SMB* being the average of the underlying two size factors formed from each of the two component double sorts, and finally FFL5F, comprising CAPM plus *SMB*, *HML*, *UMD*, and *FFL**, with these formed from three individual 3×3 double-sort procedures and size factor being the average of each constituent size factor from each double sort. MAA denotes the mean absolute alpha, in percentage terms. R_C^2 is the cross-sectional constrained R^2 , and the numbers in parentheses represent the respective empirical p-values to test the null that the explanatory ratio is zero (obtained from a bootstrap simulation). The results in Panel A are estimated across all 10 deciles for each of the 14 anomaly variables, namely, 140 (10×14) portfolios in total, while the results in Panel B are for the extreme deciles only, namely, D1, D2, and D3 and then D8, D9, D10, which results in 84 (6×14) portfolio test assets in total. Initial time-series regressions based on Huber-White heteroscedasticity-robust standard errors.

	CAPM	FF3F	Carhart 4F	ILLIQ2F	FFL2F	FFL3F	FFL4F	FFL5F
Panel A: overall sample								
MAA, %	0.529	0.198	0.201	0.360	0.522	0.181	0.198	0.200
R_C^2 [p-value]	0.049 [0.00]	0.497 [0.00]	0.455 [0.00]	0.022 [0.04]	0.060 [0.00]	0.539 [0.00]	0.342 [0.00]	0.501 [0.00]
Panel B: extreme deciles								
MAA, %	0.561	0.241	0.240	0.384	0.552	0.221	0.236	0.210
R_C^2 [p-value]	0.039 [0.02]	0.473 [0.00]	0.437 [0.00]	-0.047 [0.44]	0.046 [0.02]	0.531 [0.000]	0.517 [0.00]	0.488 [0.00]

Table 7 Spanning regressions

This table presents the results for the US S&P1500 universe for the period 1985:01 to 2023:12. The results for the maximum Sharpe ratio, $Sh^2(f)$, are reported in Panel A for each of the eight multi-factor asset pricing models. These are the FF3F [CAPM plus *SMB*, *HML*], Carhart 4F [FF3F plus *UMD*], ILLIQ2F and FFL2F followed by FFL3F [CAPM plus *SMB*, *FFL**], FFL4F [CAPM plus *SMB*, *HML*, *FFL**] and FFL5F [CAPM plus *SMB*, *HML*, *UMD*, *FFL**]. Spanning regressions and marginal contributions to each asset pricing model's maximum Sharpe ratio, $Sh^2(f)$, for each individual factor regression for factors contained within each of the seven models are reported in Panels B to H respectively. Regression intercepts are termed "a," the market excess returns are " $R_m - R_f$," then for all multifactor models with three factors or over, *SMB*, *HML*, *UMD*, and *FFL** are all formed through a 3×3 double-sort procedure first on size then the factor concerned. However, only in the two factor models ILLIQ2F and FFL2F are the *ILLIQ* and *FFL* factors obtained through returns differences between highest (lowest) and lowest (highest) sorted extreme decile portfolios. The table shows intercepts a, t-statistics for the intercepts t(a), slopes, R^2 , and residual standard errors s(e) from spanning regressions of each of the factors of a model on the model's other four or five factors. The table also shows $Sh^2(f)$ and each factor's marginal contribution to a model's $Sh^2(f)$, that is, $a^2/s^2(e)$. Standard errors are Huber-White heteroscedasticity-robust.

Panel A: Max. Sharpe ratios

FF3F	Carhart 4F	ILLIQ2F	FFL2F	FFL3F	FFL4F	FFL5F	FFL5F
0.447	0.485	0.243	0.111	0.438	0.455	0.489	0.496

Panel B: Spanning regressions FF3F

LHS	a	R _m -R _f	SMB	HML	t(a)	R ²	s(e)	Sh ² (f)	a ² /s ² (e)
R _m -R _f	0.011	-- --	-0.251	-0.411	2.97	0.057	0.075	0.447	0.021
SMB	0.013	-0.043		0.170	9.35	0.056	0.031	0.447	0.186
HML	-0.003	-0.089	0.210	-- --	-1.91	0.081	0.035	0.447	0.008

Panel C: Spanning regressions Carhart 4F

LHS	a	R _m -R _f	SMB	HML	UMD	t(a)	R ²	s(e)	Sh ² (f)	a ² /s ² (e)
R _m -R _f	0.013	-- --	-0.254	-0.542	-0.295	3.43	0.082	0.074	0.485	0.030
SMB	0.013	-0.045	-- --	0.160	-0.019	9.02	0.057	0.031	0.485	0.189
HML	-0.001	-0.103	0.170		-0.302	-0.30	0.214	0.032	0.485	0.001
UMD	0.007	-0.089	-0.033	-0.479	-- --	4.02	0.156	0.041	0.485	0.035

Panel D: Spanning regressions ILLIQ2F

LHS	a	R _m -R _f	ILLIQ	t(a)	R ²	s(e)	Sh ² (f)	a ² /s ² (e)
R _m -R _f	0.009	-- --	0.071	2.55	0.004	0.077	0.243	0.015
ILLIQ	-0.015	0.058	-- --	-4.66	0.004	0.070	0.243	0.047

Panel E: Spanning regressions FFL2F

LHS	a	R _m -R _f	FFL	t(a)	R ²	s(e)	Sh ² (f)	a ² /s ² (e)
R _m -R _f	0.008	-- --	0.271	2.22	0.019	0.077	0.111	0.011
FFL	0.001	0.071	-- --	0.56	0.019	0.039	0.111	0.001

Panel F: Spanning regressions FFL3F

LHS	a	$R_m - R_f$	SMB	FFL*	t(a)	R ²	s(e)	Sh ² (f)	a ² /s ² (e)
$R_m - R_f$	0.013	-- --	-0.404	1.008	3.61	0.084	0.074	0.438	0.033
SMB	0.013	-0.074	-- --	0.219	9.19	0.037	0.032	0.438	0.183
FFL*	-0.001	0.063	0.075	-- --	-2.57	0.071	0.018	0.438	0.063

Panel G: Spanning regressions FFL4F

LHS	a	$R_m - R_f$	SMB	HML	FFL*	t(a)	R ²	s(e)	Sh ² (f)	a ² /s ² (e)
$R_m - R_f$	0.012	-- --	-0.313	-0.362	0.942	3.24	0.112	0.073	0.455	0.026
SMB	0.013	-0.057	-- --	0.173	0.228	9.47	0.074	0.031	0.455	0.193
HML	-0.003	-0.083	0.218		-0.096	-2.01	0.083	0.035	0.455	0.009
FFL*	-0.001	0.061	0.081	-0.027	-- --	-2.66	0.074	0.018	0.455	0.072

Panel H: Spanning regressions FFL5F

LHS	a	$R_m - R_f$	SMB	HML	UMD	FFL*	t(a)	R ²	s(e)	Sh ² (f)	a ² /s ² (e)
$R_m - R_f$	0.013	-- --	-0.315	-0.489	-0.285	0.928	3.68	0.135	0.072	0.489	0.036
SMB	0.013	-0.058	-- --	0.162	-0.021	0.228	9.14	0.074	0.031	0.489	0.195
HML	-0.001	-0.098	0.175	-- --	-0.301	-0.072	-0.38	0.215	0.032	0.489	0.000
UMD	0.007	-0.091	-0.036	-0.478	-- --	0.036	4.04	0.156	0.041	0.489	0.036
FFL*	-0.001	0.062	0.081	-0.023	0.007	-- --	-2.70	0.074	0.018	0.489	0.081

Appendix. Variable definitions

All data are from Refinitiv Datastream and Worldscope (accessed through Datastream)

Variable	Definition	Datastream Mnemonic
Panel A: Essential variables		
Price	This is the adjusted default official daily closing price. It is denominated in primary units of local currency. Prices are generally based on the “last trade” or an official price fixing. The “current” prices taken at the close of the market are stored each day. These stored prices are adjusted for subsequent capital actions, and this adjusted figure then becomes the default price available.	P
Number of shares	This is the total number of ordinary shares issued and outstanding that represent the capital of the company. The data type is expressed in thousands.	NOSH
Volume	This shows the number of shares traded for a stock on a particular day. The data type is reported in thousands. Daily and non-daily figures are adjusted for capital events. However, if a capital event occurs in the latest period of a non-daily request, then the volume for that particular period is only retrieved as unadjusted.	VO
Risk-free rate	Defined as the US\$ 1-month deposit middle rate [ECUSD1M] in Datastream. This is expressed annually; thus, it is transformed into a monthly equivalent.	ECUSD1M
Panel B: Anomaly variables		
Free float	The percentage of total shares in an issue available to ordinary investors (NOSHFF). This means the total number of shares (NOSH) less the strategic holdings (NOSHST). In general, only holdings of 5% or more are counted as strategic. Strategic ownership data are collected by the Refinitiv Reuters Ownership team. The data are derived from 11 primary sources, including SEC filings (such as Schedule 13D and Form 13FD) and the UK Register, as well as annual reports, interim reports, stock exchanges, official regulatory bodies, third-party vendors, company websites, approved news sources, and direct contact with company investor relations departments. Ownership updates were obtained at the end of each month before August 2009; after this date, values are updated on the 10th and 30th of each month. Strategic holdings are defined as the sum of the following shareholding categories: (1) Government: State (government) or state (government) institution (NOSHGV) (2) Cross Holdings: Holdings by one company in another (NOSHCO) (3) Pension Fund: Pension funds or Endowment funds (NOSHPPF) (4) Investment Co.: Investment banks or institutions seeking a long-term return. Note that holdings by hedge funds are not included (NOSHIC) (5) Employees: Employees, or by those with a substantial position in a company that provides significant voting power at an annual general meeting (typically family members) (NOSHMEM) (6) Other Holdings: Entities outside one of the above categories (NOSHOF) (7) Foreign Block Holders: Holdings by an institution domiciled in a country other than that of the issuer (NOSHFR).	NOSHFF
$\beta_{\text{pre-rank}}$	Pre-ranking Beta is measured at the end of each June; it is the sum of the slopes from the regression of monthly returns on the current and first lag of monthly market returns. The regression uses the preceding 60 months of returns. It is calculated as the rolling 60-month (5 years $\times$ 12 months) beta from rolling window single-factor capital asset pricing model regression using the S&P 1500 aggregate value-weighted benchmark for each individual stock.	-- --

MCAP	First, the total number of shares issued and outstanding [NOSH] of each firm is rebased from its originally reported in thousands in Datastream to millions. Next, this value [NOSH] is multiplied by the closing price [P]. All data are daily frequency.	-- --
B/M	The ratio of the book value of equity to the market value of equity. Book equity is the total assets for the last fiscal year-end in calendar year $t - 1$, minus liabilities, plus balance sheet deferred taxes and investment tax credit if available, minus the preferred stock liquidating value if available, or the redemption value if available, or the carrying value, adjusted for net share issuance from the fiscal year-end to the end of December of $t - 1$. Market equity (market cap) is the price times shares outstanding at the end of December of $t - 1$. In Datastream, it is reported as the market-to-book value (Worldscope item 03501), implying the inverse of this item should be taken.	BTMV
Volatility	Volatility is measured across the preceding 12 months in terms of the average daily stock price return variance, in terms of daily closing price returns [P]. Stock prices are denominated in local currency units.	-- --
Asset growth	This is a Datastream-specific term for the change in the total assets of the firm between the total assets reported two years before to one year before, $(\text{Total Assets } (-2) - \text{Total Assets } (-1)) / \text{Total Assets } (-2)$, and is expressed in percentage change terms.	WC08625
Accrual	The change in operating working capital per split-adjusted share from $t - 2$ to $t - 1$ divided by book equity per split-adjusted share at $t - 1$. Operating working capital is current assets minus cash and short-term investments minus current liabilities plus debt in current liabilities. Accrual is defined by $(\text{Current asset [WC02201]} + \text{Short-term investment [WC02001]}) - (\text{Current liability [WC03101]} + \text{Short-term debt and current portion of long-term debt [WC03051]})$, while it is scaled by book equity, defined as total shareholder equity [WC03995].	-- --
OP	Operating profit is measured with accounting data for the fiscal year ending in year $t - 1$ and is revenues minus the cost of goods sold, minus selling, general, and administrative expenses, minus interest expenses, all divided by book equity. This is defined as operating income scaled by total shareholder equity, or book equity $(\text{operating income [WC01250]} / \text{Total shareholder equity [WC03995]})$.	-- --
DY	Dividend yield is defined as the total dividends paid out from July of year $t - 1$ to June of year t divided by the market equity (from CRSP) at the end of June of year t . This is a Worldscope data item ["DY"].	DY
P/CF	Price-to-cash-flow ratio is defined as the stock price to cash flow per share. It is defined as the market equity at the end of December of $t - 1$ divided by the cash flows for the fiscal year ending in calendar year $t - 1$. Cash flows are income before extraordinary items plus depreciation. Firms with nonpositive cash flows are excluded.	WC09604
NSI	Net issues (2 year $- 1$ year) is defined as follows: At the end of June of year t , we measure net stock issues as the natural log of the ratio of the split-adjusted shares outstanding at the fiscal year ending in calendar year $t - 1$ to the split-adjusted shares outstanding at the fiscal year ending in $t - 2$. This is defined as the natural logarithm of the difference in the total shares issued and outstanding for the firm two years before to one year before, namely, $\text{Ln}(\text{Total shares outstanding [NOSH]}(-1)) - \text{Ln}(\text{Total shares outstanding [NOSH]}(-2))$.	-- --
P/E	Price-to-earnings ratio for valuing a company that measures its current share price relative to its per-share earnings.	PE
CAPX Growth	This is calculated from the change in the capital expenditure (CAPX) of the firm between the CAPX reported three years [WC04601] before to one year before [WC04601], $(\text{CAPX [WC04601]} (-3) - \text{CAPX [WC04601]} (-1)) / \text{CAPX [WC04601]} (-3)$ and is expressed in percentage change terms. Capital Expenditure (CAPX) represent the funds used to acquire	-- --

	fixed assets other than those associated with acquisitions. It includes but is not restricted to additions to property, plant and equipment, investments in machinery and equipment but net of disposal.	
Sales Growth	This is defined as $((\text{Current Year's Net Sales or Revenues} / \text{Net Sales or Revenues six years ago, reduced to a compound annual rate}) - 1) * 100$.	WC08635
LIU	Our measure of illiquidity is the Liu (2006) multidimensional “trading speed” metric, which is the standardized turnover-adjusted number of zero daily trading volumes over the previous 12 months: $LM_{12} = (\text{No. daily zero volumes in prior 12 months}) + \left(\frac{\frac{1}{12 \text{ month turnover}}}{\text{Deflator}} \right) \times \left(\frac{21 \text{ days} * 12 \text{ months}}{\text{NoTD}} \right),$ where the 12-month turnover is calculated as the sum of the daily turnover over the previous 12 months, daily turnover is the ratio of the number of shares traded on a day [VO] to the number of shares outstanding at the end of the day [NOSH], NoTD is the total number of trading days in the market over the previous 12 months [derived from either “VO” or “P” data], and Deflator is chosen such that $0 < \left(\frac{\frac{1}{12 \text{ month turnover}}}{\text{Deflator}} \right) < 1$ for all sample stocks. Following Liu (2006), a deflator of 480,000 is used in constructing estimates for the Liu metric, which is selected based on $0 < \text{Deflator} < 1$. Given the turnover adjustment (the second term in brackets in the first/top expression), two stocks with the same integer number of zero daily trading volumes can be distinguished. The one with the larger turnover is more liquid. Thus, the turnover adjustment acts as a tiebreaker when sorting stocks based on the number of zero daily trading volumes over the previous X months. Because the number of trading days can vary from 15 to 23, multiplication by the factor $(21 \times 12 / \text{NoTD})$ standardizes the number of trading days in a month to 21, which makes the liquidity measure comparable over time.	-- --

about ECGI

The European Corporate Governance Institute has been established to improve *corporate governance through fostering independent scientific research and related activities*.

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

ECGI Working Paper Series in Finance

Editorial Board

Editor	Nadya Malenko, Professor of Finance, Boston College
Consulting Editors	Renée Adams, Professor of Finance, University of Oxford Franklin Allen, Nippon Life Professor of Finance, Professor of Economics, The Wharton School of the University of Pennsylvania Julian Franks, Professor of Finance, London Business School Mireia Giné, Associate Professor, IESE Business School Marco Pagano, Professor of Economics, Facoltà di Economia Università di Napoli Federico II
Editorial Assistant	Asif Malik, Working Paper Series Manager

Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute's Web-site (www.ecgi.global/content/working-papers) or SSRN:

Finance Paper Series	http://www.ssrn.com/link/ECGI-Fin.html
Law Paper Series	http://www.ssrn.com/link/ECGI-Law.html