Non-Standard Errors in Portfolio Sorts

Dominik Walter*   Rüdiger Weber‡   Patrick Weiss†

This version: April 24, 2023

Abstract

We systematically study the variation in returns induced by varying 14 methodological decisions in portfolio sorts. These *non-standard errors* range between 0.14 and 0.39 percent per month and are larger than standard errors. However, for most sorting variables, mean return differentials and alphas are pervasively positive, statistically significant, and increase monotonically. Decisions such as excluding firms with negative earnings or the information time lag have an impact comparable to size-related ones. Non-standard errors are countercyclical, raising concerns about non-classical measurement error in predictive regressions. Using our publicly available code to report distributions of estimated premia provides an easy remedy.

Keywords: Non-standard errors, portfolio sorts, data mining, p-hacking, risk factors, anomalies

JEL: C58, G10, G11, G12, G14

* Vienna Graduate School of Finance (VGSF), Welthandelsplatz 1, 1020 Vienna, Austria. E-mail: dominik.walter@vgsf.ac.at

‡ WU Vienna University of Economics and Business and VGSF, Welthandelsplatz 1, 1020 Vienna, Austria. E-mail: ruweber@wu.ac.at

† Reykjavik University, Menntavegur 1, 101 Reykjavik, Iceland and WU Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria. E-mail: patrickw@ru.is

We thank Guillaume Coqueret, Lukas Handler, Mathias Hasler, Robert Korajczyk, Parisa Mofakham, Matthias Molnar, Kevin Schneider, Matthias Reiner, Akos Török, Stefan Voigt, Marko Weber, and Josef Zechner as well as participants at Australasian 2022, AWG 2022, PFMC 2022, and seminar participants at the University of Vienna, for their comments, insights, and discussion. We provide an Internet Appendix via Github. Dominik Walter gratefully acknowledges financial support by the FWF (Austrian Science Fund).

1 Introduction

Portfolio sorts are essentially a way to estimate a nonlinear mapping from stock characteristics to expected returns. Altering seemingly innocuous choices made in a portfolio sort – e.g., the exclusion or inclusion of penny stocks or specific filters on stock characteristics – is therefore informative of the stability of the estimated functional relation but also of the underlying drivers of portfolio returns. Making different methodological choices can lead to vastly different conclusions regarding the stability of said functional relations. For instance, while Jensen, Kelly, and Pedersen (2021) successfully replicate roughly 82% of asset pricing factors in their sample, Hou, Xue, and Zhang (2020) can only replicate 35% of asset pricing factors in their sample. These differences in results are largely due to differences in methodological choices for portfolio sorts.[1]

In this paper, we systematically test how much estimated return premia vary with these methodological decisions and investigate the impact of seemingly innocuous choices in portfolio sorts. That is, we study *non-standard errors* in the spirit of Menkveld et al. (2022) in a widely used, fairly standardized procedure in asset pricing. These non-standard errors add uncertainty about the size of return premia in addition to well-understood standard errors. Specifically, we answer the following three questions:

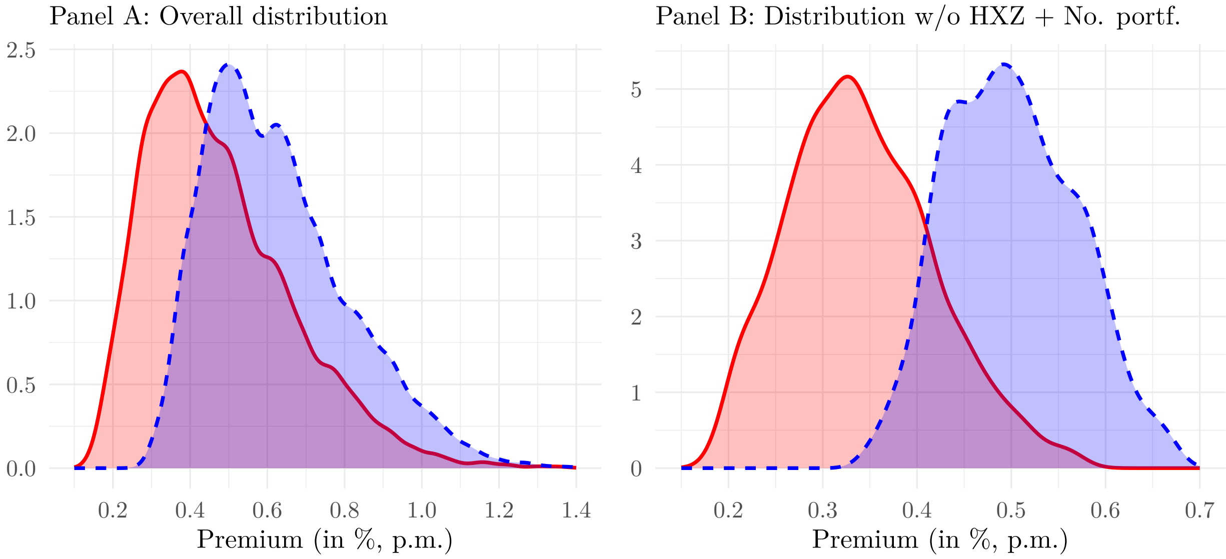
1. How large are non-standard errors in portfolio sorts?
2. Which methodological choices induce the largest variation in estimated premia?
3. What are the economic drivers of non-standard errors in portfolio sorts?

To summarize, we find that different methodological choices have a profound impact on the estimated premia that is on average larger than standard errors. 13 of the 14 decision nodes we consider have a material effect on estimated premia. Among the most impactful nodes, we do not only find expected candidates such as those discussed by Hou et al. (2020, HXZ) but also the choice to exclude firms with negative earnings, or the time period between information arrival and portfolio formation. Thirdly, we find that the size of non-standard errors varies strongly over time and is positively related to measures of volatility, economic downturns, and illiquidity.

We proceed as follows: In order to answer the first question, we analyze premium distributions generated by varying 14 methodological choices for 68 sorting variables. We show one of these distributions for the sorting variable "asset growth" (AG) in Figure 1.

[1] Hou et al. (2020) use decile portfolios based on NYSE breakpoints and value weights, whereas Jensen et al. (2021) implement tercile portfolios using the 80% largest NYSE stocks for breakpoints and capped value weights.

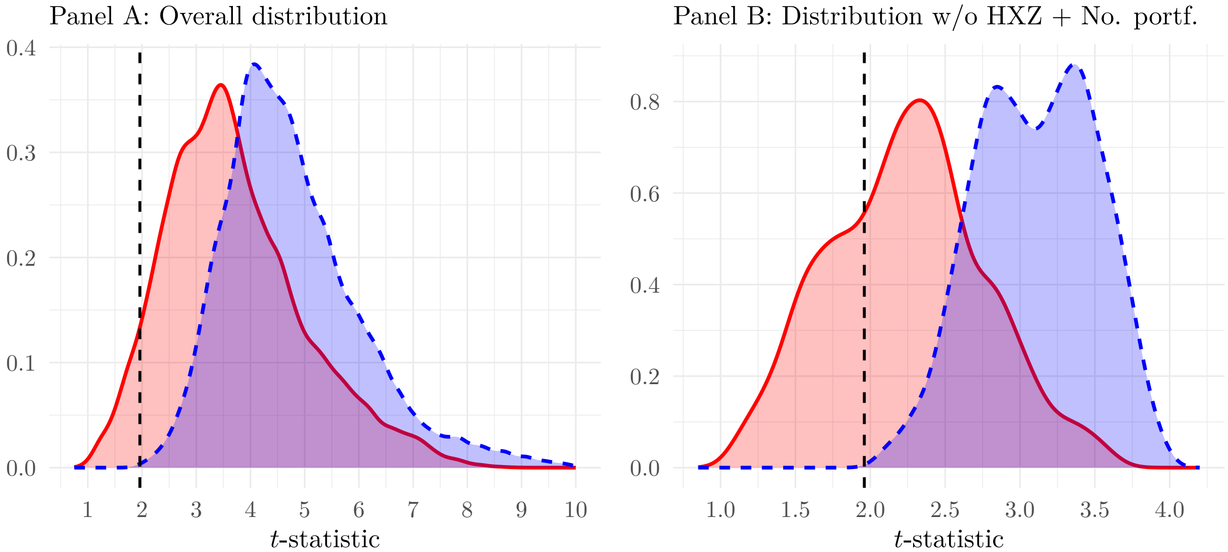
Figure 1: Non-standard errors for portfolios sorted on asset growth.
This figure shows the distribution of estimated premia (i.e., the time-series average of long-minus-short portfolio returns). Panel A shows all estimated premia implied by varying over 14 decision nodes, and Panel B only considers variation across ten decision nodes, conditional on deciles, NYSE breakpoints, and value-weighted returns from single sorts (Hou et al., 2020, HXZ). Each panel shows premia (red, solid line) and alphas of the CAPM (blue, dashed line).



The red distribution in Panel A captures variation in estimated premia across all possible decisions taken in the 14 decision nodes. Estimated premia vary widely, yielding an interquartile range of 0.26% per month, compared to an average premium across specifications of 0.48%. Controlling for the market risk factor (or other factors, for that matter) hardly reduces this variation (blue distribution). Moreover, the variation in Panel B of Figure 1 is still substantial when keeping the more obviously decisive decision nodes from HXZ constant: decile portfolios, NYSE breakpoints, value-weighted returns, and single sorts. Although methodological choices induce large variation in the estimated asset growth premia, all of these premia in Figure 1 are larger than zero, and most of the corresponding t-statistics in Figure 2 are larger than 1.96. These findings are largely representative of the remaining 67 sorting variables investigated in our paper. Strikingly, we find that premia are very robust and that the link between sorting variables and returns does not depend on specific methodological choices for the vast majority of sorting variables. Moreover, for most specifications across sorting variables, portfolio sorts generate statistically significant and monotonically increasing return spreads. Exceptions center around variables related to trading frictions and default risk where one would not expect monotonicity to begin with. Finally, differences in factor exposures do not materially affect non-standard errors.

With these findings, we do not only contribute to the growing literature on non-standard errors in finance but also to the "replication crisis" literature (see, e.g., Harvey, 2019; Jensen et al., 2021; Chen, 2022, for discussions). Rather than to check if a specific, single return differential clears ever-increasing t-hurdles, we suggest evaluating the distribution generated by varying methodological choices. Our

Figure 2: Variation in t-statistics for portfolios sorted on asset growth.
This figure shows the distribution of t-statistics of the estimated premia (i.e., the time-series average of long-minus-short portfolio returns) implied by varying over 14 methodological decision nodes. We use Newey and West (1987) standard errors with automatic lag selection as in Newey and West (1994). Panel A shows all possible specifications, and Panel B only considers variation across ten decision nodes, conditional on deciles, NYSE breakpoints, and value-weighted returns from single sorts (Hou et al., 2020, HXZ). Each panel shows premia (red, solid line) and alphas of the CAPM (blue, dashed line). A t-value of 1.96 is indicated by the vertical dashed line.



paper's evidence is reassuring because almost all considered sorting variables yield positive premia irrespective of the choices made. The evidence looks slightly less positive for t-statistics. In a little less than 50% of all cases, the t-statistic is below 1.96. The "asset growth" anomaly shown in Figures 1 and 2 is representative of both of these stylized facts. Which of these two pieces of evidence looms larger is debatable. In any case, most of the premia we investigate are remarkably stable regarding their sign and significance. Regardless, they still exhibit a wide variation that casts doubt, if not on the existence of the premium itself, then on their size, economic source, and significance.

This large variation in premia motivates us to analyze our second research question: Which methodological choices are mostly responsible for this observed variation in outcomes? To evaluate the impact of individual decision nodes, we analyze the mean absolute differences of premia which only differ in the choice for one specific decision node under investigation. In particular, we find that the number of portfolios, the weighting scheme (value vs. equal weighting), the decision to exclude stocks with negative earnings, a size filter as well as the time lag between information arrival and portfolio formation have the largest impact. Controlling for factor exposures does not materially change this result. Additionally, we find that the impact of decision nodes varies widely across sorting variables. For instance, the inclusion or exclusion of financial firms has a great impact on estimated premia for profitability variables; the frequency of rebalancing matters greatly for investment variables and the size restrictions for variables related to trading frictions. With respect to a potential "replication crisis", these findings highlight critical decision nodes where "data mining" is most likely to yield significant results. Researchers and peer reviewers may want to pay close attention to them. We also quantify the variation in estimated premia that is left once we fix specific nodes but allow for variation in all the remaining nodes. This allows us to assess the importance of non-standard errors conditional on other decision nodes. We find that non-standard errors remain large when fixing a specific choice for any of the six most decisive nodes, e.g., always using value-weighted portfolio returns.

Lastly, we analyze the third research question to shed light on the underlying economic drivers of non-standard errors in portfolio sorts. Specifically, we investigate how the time series of non-standard errors relate to various economic state variables. We find that non-standard errors in portfolio sorts are countercyclical, i.e., they are high when financial markets are volatile or illiquid, as well as in recessions. This countercyclicality is particularly strong for sorting variables belonging to the groups of profitability, size, and trading frictions. Our findings indicate that non-standard errors constitute a measurement error in estimated premia. These are correlated with economic state variables, leading to a biased coefficient estimates when predicting return premia.

What do our answers to these three research questions imply for financial economists? Since we find considerable variation induced by methodological choices, we recommend investigating the distribution of premia generated by varying over the decision nodes. The resulting distribution of premia is more informative than reporting just one premium generated by one (potentially arbitrary) specification, with robustness checks in an appendix. To encourage the adoption of this suggestion, we provide our code online at `https://github.com/patrick-weiss/PortfolioSorts_NSE`. Moreover, financial economists can also investigate which decision nodes have the largest impact on the proposed premium. This analysis is informative about the underlying drivers of the suggested premium. Lastly, our finding that variables related to the business cycle and volatility drive non-standard errors has an important implication for statistical inference. Namely, using such variables as return predictors may bias coefficients.

Our paper is related to different strands of the literature. We build on the paper by Menkveld et al. (2022), who introduce the term "non-standard errors" for the variation in estimates driven by the choices researchers make. We study a particularly important and fairly standardized instance where non-standard errors occur, namely portfolio sorts. We formally investigate variation induced by exhaustively varying decisions across 14 common nodes in a systematic way. Since it is specific to each sorting variable, we do not consider variation in the construction of sorting variables (see, e.g., Hasler, 2021). More generally, our decision nodes can be viewed as revealing a lower bound on non-standard errors, as even more nodes are conceivable. That said, we select a representative amount of nodes with sensible choices, which show the significance of non-standard errors.

Investigating non-standard errors in finance is a dynamic field. Mitton (2022) analyzes methodological variation in corporate finance regressions and finds that the selection and transformation of variables and outlier treatment are the main drivers of significance. Coqueret (2022) investigates "forking paths" in finance research (corresponding to nodes in our paper) and studies the statistical significance of portfolio strategies as an application. In line with our findings, he finds a large variation in t-statistics that can be used for p-hacking. Our interpretation is different. To us, a persistently positive premium shows that an anomaly is pervasive across stocks and evidence in favor of the existence of the anomaly from an economic point of view. That said, we also find large variation in the size of the premia.

Soebhag et al. (2023) analyze the effects of portfolio construction choices on factor models. While Soebhag et al. are interested in the performance of selected factor models, we provide a comprehensive analysis of non-standard errors in portfolio sorts and their economic interpretation. We also differ in our conclusion. Rather than suggesting fixing decisions identified by Hou et al. (2020), we advocate embracing non-standard errors by studying distributions of premia. This has the advantage of allowing for return differentials generated by a variety of underlying economic rationales, including mispricing and reasons related to the market microstructure. In contrast, e.g., simply removing all small stocks would take out important data points for studying mispricing.

We also contribute to the literature in empirical asset pricing discussing p-hacking and data mining. Prominently, Harvey (2017) discusses these issues and proposes a higher p-value threshold of three for subsequent return anomalies. Moreover, the vast number of asset pricing return anomalies have received extensive scrutiny (in, e.g., Cochrane, 2011; Harvey et al., 2016; Linnainmaa and Roberts, 2018; Chordia et al., 2020; Feng et al., 2020, among others). McLean and Pontiff (2016) show the (lack of) robustness of anomalies after their publication. In contrast, we do not consider the time frame to be at the discretion of researchers in this study. Hou et al. (2020) advocate holding certain decision nodes constant for future publications and show that many published anomalies fail significance tests in their setting. Moreover, Hasler (2022) shows that published articles consistently report premia in the right tail of the respective distribution. This is related to the *file drawer problem* of Rosenthal (1979), discussed in the specific context of financial economics by Kim and Ji (2015) and Morey and Yadav (2018). Reporting distributions of premia that embrace non-standard errors is a transparent and robust way to address data mining in portfolio sorts.

2 Data and methodology

In this section, we provide all sources of data and the 68 sorting variables constructed from this data. Then, we explain the portfolio sorting procedure alongside the 14 methodological decision nodes. Finally, we introduce the tools used to analyze the methodology-induced variation from these sorts.

In recent years, reproducibility and the need for sharing code have become central discussion points in the academic profession. Therefore, and to encourage the implementation of our approach, we share our code publicly at `https://github.com/patrick-weiss/PortfolioSorts_NSE`. Moreover, we refer the reader to Scheuch, Voigt, and Weiss (2023) for insights into the code design (see also `www.tidy-finance.org`).

2.1 Data

Our analysis relies on standard data used in most empirical asset pricing studies from 1968 until 2021. In particular, we use data on prices, returns, the number of shares outstanding, industry classifications, and trading volume of U.S. common stocks traded at NYSE, AMEX, and NASDAQ. We use delisting returns according to Shumway (1997) and set missing delisting returns with delisting codes 400–591 to -30%. Accounting data are from Compustat's North America Fundamentals Annual file. In all our analyses, we only consider observations with a valid primary link between CRSP's permno and Compustat's gvkey reported in the respective linking table.

Moreover, we obtain the return time series for the Fama and French (2015) factors from Kenneth French's website and data for the Hou et al. (2021) five-factor model from Lu Zhang's website. For our economic indicators, we use the CBOE volatility index from CRSP, the NBER recession indicator from the Federal Reserve Bank of St. Louis, the sentiment index from Baker and Wurgler (2006), and the liquidity index from Pástor and Stambaugh (2003).[2]

2.2 Sorting variables

We investigate 68 sorting variables suggested by previous studies to predict the cross-section of equity returns. They cover a wide range of suggested underlying economic mechanisms. We provide the complete list in Table 1. To facilitate comparisons, we follow Hou et al. (2020) and assign all sorting variables to one out of eight groups: Financing, intangibles, investment, momentum, profitability, size, trading frictions, and valuation. Similar to Chen and Zimmermann (2022), we distinguish between sorting variables that have been suggested to be significant predictors of the cross-section of expected returns and those that were insignificant in the original paper. Sorting variables that were never found

[2] We thank Kenneth French, Lu Zhang, Jeffery Wurgler, and Robert F. Stambaugh for providing these data.

to have a significant relation to the cross-section of stock returns in the original paper are marked with an asterisk (*) in all tables and figures. Note that we denote a sorting variable to be insignificant in the original reference paper if either Chen and Zimmermann (2022) or Jensen et al. (2021) classify them as insignificant in the original reference paper. Moreover, the construction of these sorting variables closely follows Hou et al. (2020), and we document all details in Appendix A.

We acknowledge that the construction of sorting variables also requires methodological choices, such as the treatment of missing data or the choice of estimation windows. Although these decisions also induce variation in results (see, e.g., Hasler, 2021), our focus is on the methodological decisions when mapping these sorting variables into expected returns, and we treat the definition of sorting variables as given.

2.3 Decision nodes

Researchers and practitioners have to make several decisions when implementing portfolio sorts. We follow the operations research literature (see, e.g., Kamiński et al., 2018), and label each decision as a decision *node*, which has a fixed set of possible *choices*. A collection of choices made at each of the nodes is called a *specification*. In order to assess the impact of a specific decision node, we compare the specifications from all paths conditional on a specific choice of this decision node. Therefore, we introduce the term *branch*, which subsumes all possible specifications conditional on a specific choice in the decision node under investigation.

We consider 14 common methodological decisions depicted as a flowchart in Figure 3. These decisions can generally be grouped into sample construction and portfolio construction nodes. Following this distinction, we present the seven sample construction nodes in Section 2.3.1 and the seven portfolio construction nodes in Section 2.3.2. We focus on the most common set of decision nodes in published, peer-reviewed articles and only consider choices found in the same articles. However, our choice of decision nodes analyzed in this paper can induce another layer of variation in non-standard errors (i.e., which one may call *non-standard errors of non-standard errors*). Therefore, we analyze how robust our findings are to holding specific decisions constant.

Table 1: List of 68 sorting variables.

We document the group, data frequency, abbreviation, description, publishing authors, data availability and significance in the original reference paper for all 68 sorting variables. An asterisk (*) indicates that the sorting variable is not significantly related to the cross-section of stock returns in the original reference paper. Note that we denote a sorting variable to be insignificant in the original reference paper if either Chen and Zimmermann (2022) or Jensen et al. (2021) classify them as insignificant in the original reference paper.

Group	Data freq.	Abb.	Description	Publication	Data availability
Financing	yearly	CDI	Composite debt issuance	Lyandres et al. (2008)	01.1968 - 12.2021
	monthly	CSI	Composite share issuance	Daniel and Titman (2006)	01.1968 - 12.2021
	yearly	DBE	Change in common equity	Richardson et al. (2005)	01.1968 - 12.2021
	yearly	DCOL	Change in current operating liabilities	Richardson et al. (2005)	01.1968 - 12.2021
	yearly	DFNL	Change in financial liabilities	Richardson et al. (2005)	01.1968 - 12.2021
	yearly	NDF	Net debt financing	Bradshaw et al. (2006)	01.1972 - 12.2021
	yearly	NEF	Net equity financing	Bradshaw et al. (2006)	01.1972 - 12.2021
	yearly	NXF	Net external financing	Bradshaw et al. (2006)	01.1972 - 12.2021
Intangibles	yearly	ADM	Advertisement expenses to market equity	Chan et al. (2001)	01.1973 - 12.2021
	quarterly	CFV	Cash-flow volatility	Huang (2009)	01.1978 - 12.2021
	yearly	EPRD*	Earnings' predictability	Francis et al. (2004)	01.1968 - 12.2021
	yearly	HR	Hiring rate	Belo et al. (2014)	01.1968 - 12.2021
	yearly	KZI*	Kaplan and Zingales index for financing constraints	Lamont et al. (2001)	01.1968 - 12.2021
	yearly	LFE*	Labor force efficiency	Abarbanell and Bushee (1998)	01.1968 - 12.2021
	yearly	OL	Operating leverage	Novy-Marx (2011)	01.1968 - 12.2021
	yearly	RDM	R&D expenses to market equity	Chan et al. (2001)	01.1976 - 12.2021
	yearly	RER	Real-estate ratio	Tuzel (2010)	01.1970 - 12.2021
	yearly	TAN*	Tangibility	Hahn and Lee (2009)	01.1968 - 12.2021
	yearly	WW*	Whited and Wu index for financing constraints	Whited and Wu (2006)	01.1968 - 12.2021
Investment	yearly	ACI	Abnormal corporate investment	Titman et al. (2004)	01.1968 - 12.2021
	yearly	AG	Asset growth	Cooper et al. (2008)	01.1968 - 12.2021
	yearly	DNOA	Change in net operating assets	Hirshleifer et al. (2004)	01.1968 - 12.2021
	yearly	DPIA	Change in property, plant, and equip. to assets	Lyandres et al. (2008)	01.1968 - 12.2021
	yearly	DWC	Change in net non-cash working capital	Richardson et al. (2005)	01.1968 - 12.2021

Continued on next page

Table 1: List of 68 sorting variables.

	yearly	IG	Investment growth	Xing (2008)	01.1968 - 12.2021
	yearly	DINV	Inventory changes	Thomas and Zhang (2002)	01.1968 - 12.2021
	yearly	NOA	Net operating assets	Hirshleifer et al. (2004)	01.1968 - 12.2021
	yearly	OA	Operating accruals	Sloan (1996)	01.1968 - 12.2021
	yearly	PTA	Percent total accruals	Hafzalla et al. (2011)	01.1968 - 12.2021
Momentum	monthly	ABR	Abnormal returns around earnings' announcements	Chan et al. (1996)	01.1972 - 12.2021
	monthly	MOM	Return momentum (11-month formation period)	Fama and French (1996)	01.1968 - 12.2021
	monthly	RMOM	Residual momentum (11-month formation period)	Blitz et al. (2011)	01.1968 - 12.2021
	quarterly	RS	Revenue surprise	Jegadeesh and Livnat (2006)	01.1972 - 12.2021
	quarterly	SUE	Standardized unexpected earnings	Foster et al. (1984)	01.1972 - 12.2021
	quarterly	TES	Tax expense surprise	Thomas and Zhang (2011)	01.1976 - 12.2021
	monthly	52W	52-week high	George and Hwang (2004)	01.1968 - 12.2021
Profitability	yearly	ATO*	Asset turnover	Soliman (2008)	01.1968 - 12.2021
	yearly	BL*	Book leverage	Fama and French (1992)	01.1968 - 12.2021
	yearly	CBOP	Cash-based operating profitability	Ball et al. (2016)	01.1968 - 12.2021
	yearly	CTO*	Capital turnover	Haugen and Baker (1996)	01.1968 - 12.2021
	yearly	GPA	Gross profits to assets	Novy-Marx (2013)	01.1968 - 12.2021
	yearly	O	Ohlson's O-score	Ohlson (1980), Dichev (1998)	01.1968 - 12.2021
	yearly	OPE	Operating profits to book equity	Fama and French (2015)	01.1968 - 12.2021
	quarterly	ROA	Return on assets	Balakrishnan et al. (2010)	01.1972 - 12.2021
	quarterly	ROE	Return on equity	Hou et al. (2014)	01.1972 - 12.2021
	yearly	TBI	Taxable income to book income	Lev and Nissim (2004)	01.1968 - 12.2021
	yearly	Z*	Altman's Z-score	Dichev (1998), Altman (1968)	01.1968 - 12.2021
Size	monthly	ME	The logarithm of market equity in U.S. Dollar	Banz (1981)	01.1968 - 12.2021

Continued on next page

Table 1: List of 68 sorting variables.

Trading	monthly	AMI	Amihud illiquidity measure	Amihud (2002)	01.1968 - 12.2021
frictions	monthly	BETA	Beta relative to the market	Fama and MacBeth (1973)	01.1968 - 12.2021
	monthly	BFP	Frazzini and Pedersen beta	Frazzini and Pedersen (2014)	01.1968 - 12.2021
	monthly	DTV	Dollar trading volume	Brennan et al. (1998)	01.1968 - 12.2021
	monthly	ISKEW	Idiosyncratic skewness	Bali et al. (2016)	01.1968 - 12.2021
	monthly	IVOL	Idiosyncratic volatility	Ang et al. (2006)	01.1968 - 12.2021
	monthly	MDR	Maximum daily return	Bali et al. (2011)	01.1968 - 12.2021
	monthly	SREV	Short-term reversal	Jegadeesh (1990)	01.1968 - 12.2021
	monthly	TUR	Share turnover	Datar et al. (1998)	01.1968 - 12.2021
Valuation	yearly	AM*	Assets to market equity	Fama and French (1992)	01.1968 - 12.2021
	yearly	BM	Book equity to market equity	Davis et al. (2000)	01.1968 - 12.2021
	yearly	CFM	Cash flow to market equity	Lakonishok et al. (1994)	01.1968 - 12.2021
	yearly	DM	Debt to market equity	Bhandari (1988)	01.1968 - 12.2021
	yearly	EBM	Enterprise book equity to market equity	Penman et al. (2007)	01.1968 - 12.2021
	yearly	EM	Earnings to market equity	Basu (1983)	01.1968 - 12.2021
	yearly	NDM	Net debt to market equity	Penman et al. (2007)	01.1968 - 12.2021
	yearly	NPY	Net payout yield	Boudoukh et al. (2007)	01.1972 - 12.2021
	yearly	OCM	Operating cash flow to market equity	Desai et al. (2004)	01.1972 - 12.2021
	monthly	REV	Long-term reversal	De Bondt and Thaler (1985)	01.1968 - 12.2021
	yearly	SM	Sales to market equity	Barbee Jr et al. (1996)	01.1968 - 12.2021

2.3.1 Sample construction

We take the CRSP and Compustat database from WRDS for all common U.S. stocks as a natural starting point for our analysis because the same data is available to all researchers. All decisions thereafter can be considered methodological choices or decision nodes in our paper. We start by discussing several sample restrictions commonly implemented before sorting portfolios.

Size restriction. Researchers often decide to limit their sample by excluding stocks with small market capitalizations. In our analysis, we either consider all stocks or exclude stocks below the 5% or 20% thresholds based on the market capitalization of NYSE stocks in each month. We implement this filter first such that the order of the other sample construction choices does not affect our results. The size restriction might be particularly impactful because stocks below the 20%-NYSE-threshold account on average for 50% of all stocks in our sample. Therefore, the decision to exclude such stocks has a profound impact on the sample composition, although they only make up a small fraction of the overall market capitalization. This decision node is informative to which degree small (and potentially illiquid) stocks drive the relation between the sorting variable and mean returns.

Financials. The exclusion of stocks belonging to the financial sector with standard industrial classification (SIC) codes between 6000 and 6999 is another frequently considered choice variable and reveals to which extent the relation between the sorting variable and mean returns (premia) are related to the financial sector.[3] This might be particularly relevant because financial stocks have different balance sheet patterns compared to industrial firms and a larger exposure to periods of financial instability. Banks, in particular, face additional regulation but have been argued to enjoy favorable funding conditions. Firms in the financial industry make up roughly 15% of the number of stocks and the total market capitalization in our sample.

Utilities. We also study the effects of excluding stocks from the utility sector with SIC codes between 4900 and 4999.[4] Although utility stocks comprise only about 5% of our sample, they face more regulation and exposure to underlying commodities. However, the extent to which this restricts the proposed relation between a stock characteristic and stock returns remains unclear.

[3] We use standard industrial classification codes from CRSP (item SICCD). Given the differences between industry classifications from CRSP and Compustat (see, e.g., Guenther and Rosman, 1994; Kahle and Walkling, 1996), this could also be considered as another decision node.

[4] Note that some papers classify stocks with SIC codes between 4900 and 4999 as utilities, whereas other papers only consider stocks with SIC codes between 4949 and 4999 as utility stocks.

Positive book equity. Even though limited liability implies that the market value of equity cannot become negative, firms can have negative book equity values. For instance, if firms have sufficiently negative earnings or high goodwill impairments. Interestingly, Luo et al. (2021) show that the share of negative book equity stocks has increased steadily from 1% in 1980 to about 4% of all Compustat stocks in 2012. The authors also reject the common perception that all negative book equity stocks are distressed because roughly 50% of the firms simultaneously show positive earnings. In light of these findings, the decision to exclude negative book equity stocks is worthwhile to consider and is potentially important for sorts on valuation characteristics.

Positive earnings. Moreover, we analyze the exclusion of stocks which report negative earnings (Compustat item IB, i.e., income before extraordinary items). This decision node can have a large impact on the results of mapping sorting variables into mean returns since, over time, roughly 28% of stock-months have negative earnings. Investigating this decision node might help researchers to understand whether stocks with negative earnings impact return premia. This is particularly interesting because such firms may be young firms with low profitability which have been associated with low average returns (Hou et al., 2014).

Stock age restriction. Banz and Breen (1986) noted that Compustat often adds new firms with their full history of data to the database. This implies that the full history was only available to investors at a later point in time. This introduces a backfill bias to the information considered available to investors. Fama and French (1993) investigate this concern and claim that Compustat rarely adds firms with more than two years of historical data. Therefore, researchers often require at least two years of previous observations for all firms in the Compustat database.[5] We consider this decision node for two reasons: First, this node affects roughly 12% of firm-year observations. Second, it is potentially important for real-time trading based on a particular sorting variable.

Price restriction. Lastly, researchers often exclude so-called "penny stocks" with low absolute share prices. We consider three possible choices: no exclusion, excluding stocks that trade below 1$, or excluding stocks with prices below 5$. This decision node helps to understand to which degree the functional relation between the sorting variable and expected returns is driven by stocks that may be difficult to trade.

[5] We exclude observations based on their stock age measured in years listed in CRSP. Moreover, this decision node is by construction not available for sorting variables that require two (or more) years of data for their estimation: CSI, CFV, EPRD, RMOM, BETA, BFP, and REV.

2.3.2 Portfolio construction

The sample construction constitutes just one layer of decisions for portfolio sorts. Once the underlying sample is specified, one must decide how to construct the portfolios themselves. Again, we investigate seven decision nodes for portfolio construction.

Sorting variable lag. When forming portfolios, researchers need to make a judgment call on how many months they want to keep between the arrival of information for the sorting variable and portfolio formation. For annual accounting data, sufficiently long lags of at least six months are the common choice to ensure that the respective information is available to investors when forming portfolios. Nonetheless, long lags mask the short-term relation between the sorting variable and premia throughout the first months. Therefore, we investigate a lag of three months, six months, or as in Fama and French (1992) of at least six months for annual sorting variables. For sorting variables updated monthly, we implement lags of one, three or six months between the arrival of information and portfolio formation. Lastly, we investigate a lag of three or six months for sorting variables updated on a quarterly frequency, where we assume information arrives at the end of a quarter. We indicate the frequency of sorting variables in Column 2 of Table 1. This decision node might help researchers to understand whether the mapping from the sorting variable into mean returns is rather ephemeral or persistent.

Rebalancing. We consider rebalancing portfolios on a monthly and yearly basis as it is common in portfolio sorts. Whether there is an option for the frequency of rebalancing depends on the update frequency of the sorting variable and is only sensible for yearly sorting variables (see Table 1, Column 2). The effects from this decision node can be mainly attributed to the differences in fiscal year ends for distinct stocks and the delisting of stocks. Thereby, variation from this node shows how persistent the relation between sorting variables and mean returns is. Moreover, it can also indicate to which degree the underlying relation depends on transaction costs, which are high for frequent rebalancing.

Breakpoints: Quantiles (main). The number of quantiles we use for portfolio breakpoints is an obvious driver of estimated premia. From an economic perspective, it may be indicative of the degree of monotonicity in the underlying functional relation between stock characteristics and premia. The idea is that having more portfolios, i.e., using more extreme breakpoints, will naturally lead to a larger premium. Moreover, it is plausible that more extreme breakpoints might also have an impact on the effects of other decision nodes. In this paper, we consider either quintiles or deciles to determine the breakpoints in the main sorting variable.

Double sort. Researchers may be interested in how the relation between a sorting variable and mean returns holds up when adding another variable, such as size. Thus, we consider independent and dependent double sorts into portfolios sorted on each of the sorting variables on the one hand and size on the other. These double sorts essentially partition each portfolio of the sorting variable further into size portfolios. Dependent double sorts condition the breakpoints of the sorting variable on quantiles within the respective size portfolios. In contrast, independent double sorts compute the two sets of breakpoints independently of the other sorting variable. This decision node reveals whether the relation between the primary sorting variable and mean returns is driven by size. In contrast to other size-related nodes, double sorts account for a relation between the sorting variable and size across the entire distribution of market equity rather than just accounting for the smallest stocks.[6]

Breakpoints: Quantiles (secondary). Conditional on investigating double sorts, a researcher has to decide how many secondary portfolios to form. We allow for two or five secondary portfolios. From an economic perspective, the granularity of secondary breakpoints indicates to which extent the relation between the primary sorting variable and mean returns is robust or limited to extreme observations of the secondary variable (size). This becomes relevant if researchers are concerned that the primary sorting variable is closely related to other characteristics known to predict the cross-section of stock returns.

Breakpoints: Exchanges. To mitigate the impact of small stocks, breakpoints for the primary but also secondary sorting variables are sometimes based only on stocks listed on the New York Stock Exchange (NYSE). The NYSE stocks have an average market capitalization of roughly 5bn$ compared to 1bn$ for stocks listed on Nasdaq and 4bn$ for stocks listed on Amex. This choice of breakpoints allows for an interesting interpretation: If we observe significant non-standard errors alongside this decision node, the primary sorting variable is likely to be related to size.

Weighting scheme. After assigning stocks to their respective portfolio, a researcher has to decide how to aggregate individual stock returns into a portfolio return. We consider either equally weighted average returns or returns weighted with the market capitalization of the corresponding stock (i.e., value weights). This decision node shows to which degree the relation between the sorting variable and mean returns depends on returns from small stocks within portfolios.

[6] As a technical detail, we obtain the long (short) return for each sorting variable as the equally-weighted average return of all long (short) portfolios over the size dimension in the double sort. However, we still allow the researcher to choose a weighting scheme for the individual long (short) portfolios within this double sort.

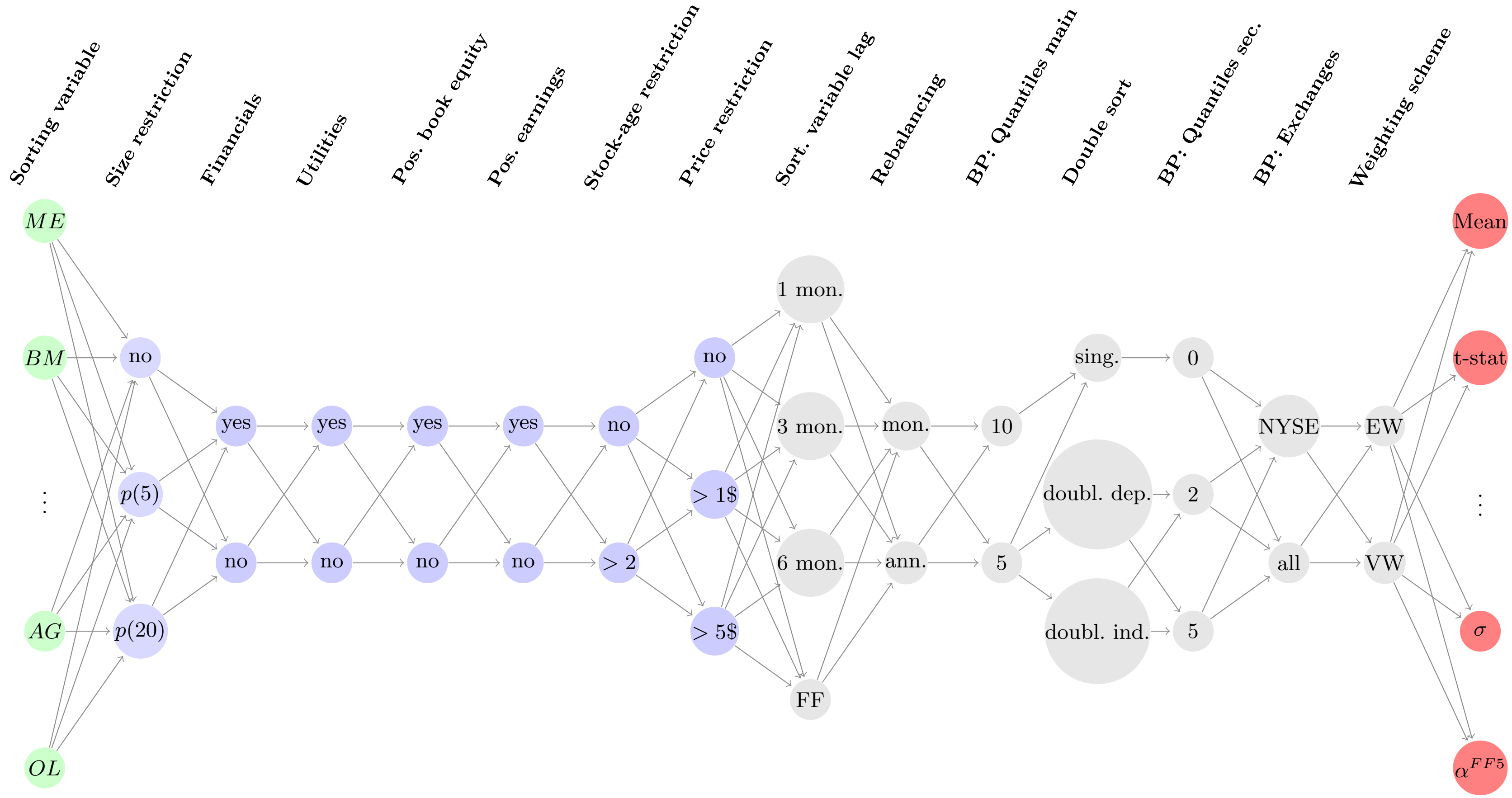
Overall, these 14 decision nodes imply 69,120 specifications for sorting variables updated on an annual basis, 34,560 specifications for variables updated monthly, and 23,040 specifications for variables updated quarterly.[7] This amount of potential specifications underlines that non-standard errors might be particularly relevant in portfolio sorts. Given that we analyze 68 sorting variables, we report the outcomes for a total of 3,738,240 specifications in the following sections.

We keep some potential nodes constant in this paper and thus do not consider them as additional decision nodes. First, we do not analyze the impact of changing the order of decision nodes in the sample construction. Any reordering would only affect our results if the size filter were applied to the subset of stocks in the sample after applying other sample restrictions because it is based on quantiles. Second, the exact definition of each sorting variable is, in principle, another methodological choice constituting another source of non-standard errors. We do not investigate various variable definitions separately, because there are no standardized procedures applicable to all sorting variables. Moreover, we also keep the sample period constant since it is not necessarily at discretion. Finally, we do not study the impact of coding errors on the outcome of portfolio sorts. Including mistakes in some code by design seems arbitrary, even though it may have a large impact. In light of these potential extensions of decision choices, our estimates can be understood as a lower bound for non-standard errors in portfolio sorts.

[7] The difference in the number of specifications between sorting variables is due to the following: First, we only consider different rebalancing frequencies for annual sorting variables. Second, depending on the frequency of information arrival, the sorting variables differ in the number of permissible choices for the sorting variable lag. Last, note that the stock age filter is not available for sorting variables requiring two or more years of data for their estimation.

Figure 3: Flowchart of decision nodes for portfolio sorts.

After constructing 68 sorting Variables (SV) we consider the paths of 14 decision nodes for portfolio sorts until the final nodes, i.e., the output. The first seven decision nodes are sample construction nodes: include large stocks dependent on market equity quantiles (all, larger then p(5) or p(20)), include financials (yes or no), include utilities (yes or no), firm-months with positive book equity (yes or no), firm-months with positive earnings (yes or no), stocks-age filters (at least two years or all), and stock prices (larger than \$1, \$5, or all). The ensuing seven decision nodes belong to the portfolio construction nodes: the lag of the sorting variables (one month, three months, six months, or a Fama and French (1992) lag), the portfolio rebalancing (monthly or annually), the number of main portfolios (5 or 10), the sorting method (single sorts, independent or dependent double sorts), the number of secondary portfolios for double sorts (2 or 5), the exchanges for breakpoints (NYSE or all), and the weighting scheme (equal- or value-weighting).



2.4 Empirical methodology

We define a specification $s = (c_1, ..., c_N)$ as a vector of choices c_n, for each decision node $n = 1, \ldots, 14$. Thus, each specification corresponds to one path in Figure 3, which, in turn, corresponds to one portfolio sort. For each sorting variable $v \in V = \{\text{CDI}, \ldots, \text{SM}\}$ and each specification $s = 1, \ldots, S$, we compute the average return differential between the two extreme portfolios (i.e., long-minus-short portfolio returns), which we refer to as *premium*[8], i.e.,

$$r_s^v = \frac{1}{T} \sum_{t=1}^{T} \left(r_{t,s}^{v,\text{Long}} - r_{t,s}^{v,\text{Short}} \right), \tag{1}$$

where $r_{t,s}^{v,\text{Long}}$ and $r_{t,s}^{v,\text{Short}}$ denote the returns of the two extreme portfolios, such that the sign is normalized to yield a positive premium in line with the direction proposed by the original paper. We obtain the extreme portfolio returns by weighting the stock returns in the portfolio according to the specification's weighting scheme.

All results shown in this paper are based on monthly returns in percent. Furthermore, we also adjust the monthly returns accounting for the exposure to the factors of the Capital Asset Pricing Model (CAPM), the Fama and French (2015) model (FF5), and the Q model with expected growth from Hou, Mo, Xue, and Zhang (2021), which we discuss in Section 3.2 and denote as Q5. Whenever we present summary statistics of the premia produced by different specifications, we take an average over the (sub-)sample of premia for each sorting variable, before averaging across sorting variables. Removing outliers does not impact our results, and we do not truncate or winsorize our samples.

Throughout the paper, we report Newey and West (1987) standard errors with automatic lag selection following Newey and West (1994). While there are different procedures to adjust standard errors, we keep this decision node constant to have comparable results. These corrected standard errors are also used for t-statistics. We aggregate t-statistics by counting the number of specifications larger than 1.96 relative to the total number of specifications within a test. We also compute average standard errors by taking the average across specifications.

We define *non-standard errors* in line with Menkveld et al. (2022) as the interquartile range of estimates across specifications. Specifically, in our case we measure *non-standard errors* for each sorting variable v as:

$$\text{NSE}^v = Q_{0.75}(r^v) - Q_{0.25}(r^v), \tag{2}$$

[8] We are deliberately loose with the term "premium". Not all considered return differentials were originally referred to as premia in the asset-pricing sense of a (risk) premium. Moreover, strictly speaking, the expected value of the long-short-portfolio return is the premium, whereas the average long-minus-short portfolio return is an estimate.

where $Q_\alpha(r^v)$ denotes the α-quantile of the distribution of r^v, which corresponds to the aggregated premia r_s^v across all specifications s. Following Menkveld et al. (2022), we evaluate the significance of *non-standard errors* for each sorting variable v by testing whether each estimated premium r_s^v is significantly different from the median premium across all specifications. Formally, we test the following hypothesis:

$$H_0 : r_s^v = Median(r^v) \qquad \forall s \in \{1, ..., S\}. \tag{3}$$

Similar to Menkveld et al. (2022), we consider *non-standard errors* for a sorting variable v to be significant if at least one of these tests rejects the null hypothesis in Equation (3). Additionally, we report the frequency of the rejected null hypotheses relative to the total number of specifications s for each sorting variable v.

We relate the size of non-standard errors to the size of standard errors. To put both quantities on equal footing, we estimate the dispersion induced by methodological choices as the standard deviation of the distribution of premia r^v across all specifications for a single sorting variable v. Then, we divide this standard deviation by the average standard error for each sorting variable v:

$$\text{Ratio}^v = \frac{\sqrt{\frac{1}{S-1}\sum_{s=1}^{S}(r_s^v - \bar{r^v})^2}}{\frac{1}{S}\sum_{s=1}^{S}\sigma_s^v}, \tag{4}$$

where $\bar{r^v}$ is the average premium across all specifications and σ_s^v is the estimated time-series standard error of r_s^v.

To investigate the economic interpretation of the return premia in our sample, we follow Patton and Timmermann (2010) and conduct a monotonicity test for all permissible pairs of portfolio returns (in untabulated results, we also test for monotonicity based only on adjacent pairs, which does not impact our conclusions). We use quintiles in these tests to avoid comparing monotonicity over a varying number of portfolios. These tests allow us to investigate how important non-standard errors in portfolio sorts are for risk-based characteristics in our sample.

We evaluate the importance of each decision node in three ways: First, we calculate mean absolute differences of the estimated premia for each of the 14 nodes. We compute these mean absolute differences for pairs of specifications (i, j) that only differ in the choice made in one specific node n. Formally, for each node n and sorting variable v, this set of specification pairs can be defined as $P_n^v = \{(i, j) \mid c_{i,m} = c_{j,m} \ \forall\ m \in \{1, \ldots, 14\} \setminus n, c_{i,n} \neq c_{j,n}\}$. At each point in time, we first calculate the absolute differences in premia between each of the pairs of specifications in P_n^v. Then we average across all different pairs P_n^v for each sorting variable, and, finally, we average across all sorting variables. Formally, the mean absolute difference at time t is given by:

$$MAD_t^n = \frac{1}{|V|} \sum_{v \in V} \left(\frac{1}{|P_n^v|} \sum_{(i,j) \in P_n^v} |r_{t,i}^v - r_{t,j}^v| \right). \tag{5}$$

We also provide results for the mean absolute difference MAD^n for each decision node n aggregated over time. To do so, we take the time-series average of MAD_t^n.

Second, the impact of a decision node n can also be evaluated by the average time-series correlation of all specification pairs P_n^v. If the time-series correlation of two estimated premia that differ only in the node under investigation is low (high), then this node has a large (small) impact on the return time series. Therefore, we calculate the average time-series correlation across all specification pairs P_n^v for all sorting variables V as follows:

$$Corr^n = \frac{1}{|V|} \sum_{v \in V} \left(\frac{1}{|P_n^v|} \sum_{(i,j) \in P_n^v} \rho_{i,j}^v \right), \tag{6}$$

where $\rho_{i,j}^v$ is the time-series correlation of the estimated premia from specification pairs in P_n^v.

Third, we follow Menkveld et al. (2022) and test whether the sample of premia that differ in the node under investigation are drawn from the same distribution. To do so, we implement a k-sample Anderson-Darling test following Scholz and Stephens (1987). Specifically, we aggregate the premia r_i^v associated with a node's branches for all sorting variables into branch-specific distributions. Subsequently, we test whether these samples are drawn from the same distribution.

3 How large are non-standard errors in portfolio sorts?

The functional relation between a sorting variable and expected returns can be estimated from portfolio sorts. However, this estimate might differ considerably, depending on specific methodological choices. To study this source of variation, we investigate the impact of 14 methodological choices on return premia estimated from portfolio sorts. Thus, we calculate the outcomes, such as premia and t-statistics, separately for each sorting variable and for all specifications generated by varying over all decision nodes. For each sorting variable, we then aggregate the premia and t-statistics from all specifications into distributions. We discuss results for long-minus-short return differentials in Section 3.1 and for factor adjusted premia in Section 3.2.

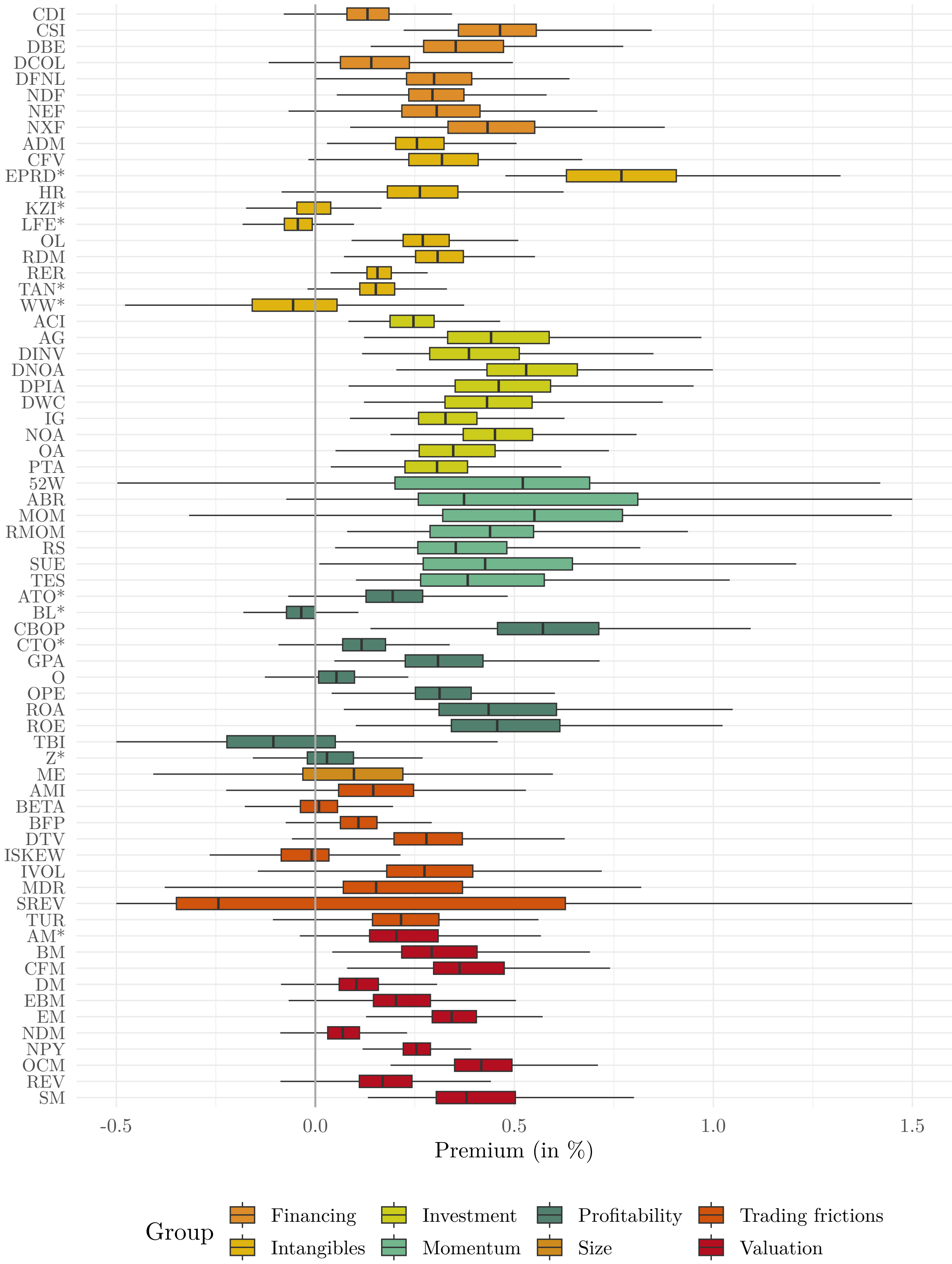
3.1 Non-standard errors in unadjusted premia

Figure 4 shows these distributions of unadjusted ("raw") premia from all possible specifications for each sorting variable. While the distributions' shapes vary widely across sorting variables, the average variation depicted by these box plots is substantial. Moreover, Figure B.1 shows that the variation in premia also translates into large variation in t-statistics.[9] Even from this graphical analysis, we can deduce that methodological choices have a profound impact on the size and significance of estimated premia. Table 2 provides an overview of statistical moments.

The average non-standard error across all sorting variables as measured by the interquartile range in Panel I of Table 2 is about 0.20% per month, which compares to an average premium of 0.29% per month. For instance, the non-standard error of 0.26% per month for the sorting variable "asset growth" (AG) suggests that the 50% of premia around the mean range from 0.35% to 0.61% per month. This corresponds to a difference of at least 3% per year for methodology-implied variation in the asset growth premium. Similar to Menkveld et al. (2022), we consider non-standard errors to be significant if at least one specification is statistically different from the median premium (Equation 3). Across all sorting variables we find strong evidence for the significance of non-standard errors. Roughly 3% of specifications on the left side and 6% on the right side of the median premium are statistically different from the median premium. Moreover, we find that the estimated standard deviations of premia even exceed the average time-series standard errors of the individual premia in Panel I of Table 2, indicating that non-standard errors are similarly important as well-understood standard errors. All these results are slightly more pronounced if we only consider predictors which were found to be significant in the original paper (Row "Orig. Sig."). All in all, Panel I shows that methodological choices in portfolio sorts induce considerable uncertainty about the size of cross-sectional premia. Their magnitude is at least as large as standard errors.

[9] To complete the analysis, we also show variation in Newey and West (1987) standard errors in the Internet Appendix, i.e., in Figure III.4.

Figure 4: Non-standard errors across sorting variables.
This figure shows the estimated premia (in %) in box plots for all sorting variables across all decision nodes. The vertical axis shows the associated sorting variable, while the color scheme connects each sorting variable to the respective group.



Although we find large non-standard errors in portfolio sorts, our findings also convey positive news for the debate on the "replication crisis" in financial economics. Roughly 90% of the long-minus-short return differentials in Figure 4 are positive and roughly 52% statistically significant at the 5% level as shown in Panel I of Table 2. This fraction of significant premia is a conservative estimate. As we discuss in Section 3.2 below, this fraction increases to 74% once we remove sorting variables, which were never significant in the original paper, and consider CAPM alphas. Moreover, the share of positive premia is at least as high as 70% for all groups of sorting variables. These findings are reassuring in the sense that the true expected return differentials are likely non-zero and that they are robust to different methodological choices. In fact, we find that literally all specifications yield positive premia for the majority of sorting variables (Column "Pos."). The distributions of premia in Figure 4 are skewed to the right and show excess kurtosis. This indicates that methodological choices have an impact beyond the interquartile range of these distributions. A conclusion from the positive skew is that researchers have considerable leeway to report strongly positive results.

Moreover, we emphasize the economic interpretations of non-standard errors and analyze the impact on the monotonicity of premia: Almost half of all specifications across sorting variables show evidence of monotonically increasing portfolio returns following the methodology of Patton and Timmermann (2010). The large fraction of monotonic premia is a sign of the stability of risk-related sorting variables. This finding is particularly remarkable: First, the test of Patton and Timmermann (2010) imposes a strong hurdle. Second, a considerable part of our sorting variables is presumably associated with mispricing, where monotonicity is not to be expected.

Apart from these general conclusions about non-standard errors across all sorting variables, we find large heterogeneity for different groups. First, we investigate differences in the estimated non-standard error. Similar to Figure 4, Table 2 also indicates that non-standard errors tend to be considerably larger for sorting variables belonging to the groups "momentum", "size", and "trading frictions". On average, we observe non-standard errors around 0.39, 0.26, and 0.26% per month for these groups. In contrast, sorting variables from the groups "intangibles" and "valuation" tend to have relatively low non-standard errors with 0.14% per month. Additionally, the fraction of specifications that are significantly different from the median is considerably smaller for these groups compared to sorts related to momentum, size, and trading frictions. Note that non-standard errors are not only heterogeneous between groups but also within groups: Non-standard errors are consistently large for all momentum-related sorting variables. On the other hand, the non-standard error in the trading frictions and intangibles group varies more between the variables. For instance, the "short-term reversal" (SREV) has the highest non-standard error overall, but two estimates of "beta" (BETA and BFP) have below-average non-standard errors in the groups of trading frictions.

Second, there are also large differences between groups of variables when we consider the share of significant and monotonic premia. Sorting variables belonging to the groups of financing, investment, and momentum have the largest share of significant premia (75%, 96%, and 81%, respectively). Additionally, they also have the largest share of monotonic premia (49%, 67%, and 70%, respectively). Not surprisingly, these robust premia typically have strong theoretical foundations. Most strikingly, investment, profitability, and valuation sorts which all build on q-theory (see, e.g., Hou et al., 2014) yield positive premia in virtually all specifications.[10] In contrast, the share of significant premia is remarkably low for sorting variables belonging to size and trading frictions (8% and 16%, respectively). What is more, only 17% and 13% of premia from the groups of size and trading frictions show a monotonic relation to mean returns. This is not surprising, given that premia for these sorting variables are expected to be driven by the most extreme stocks.

As already mentioned in the introduction, Hou et al. (2020) emphasize the relevance of size-related decision nodes for premia. Therefore, we repeat our analysis conditional on using single sorts, decile portfolios, NYSE breakpoints, and value-weighted portfolios. Although the average non-standard error is reduced to 0.14% per month in Table IV.1 in Appendix IV, it is still large and statistically significant. This implies that variation in estimated premia induced by methodological choices is by no means restricted to the choice nodes suggested by Hou et al. (2020).

Summing up, there are two main takeaways from this section. First, methodological choices induce statistically significant non-standard errors that differ between groups of sorting variables but generally produce premium distributions that are right-skewed with excess kurtosis. This finding points to the necessity of investigating non-standard errors for return premia. Secondly, the presence of sizable non-standard errors still leaves many sorting variables' premia pervasively positive and significant. This alleviates concerns of the "replication crisis" literature and confirms the existence of most premia in our sample.

[10] Remarkable exceptions are leverage and default risk variables which are assigned to the profitability category in Hou et al. (2020) but have different theoretical underpinnings.

Table 2: Non-standard errors across sorting variables.
This table shows summary statistics across all specifications for individual sorting variables in panels grouped by categories. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level. Finally, the overall means of the statistics across all sorting variables are reported in the last panel. An asterisk (*) next to the name of the sorting variable (SV) indicates that it is not significantly related to the cross-section of stock returns in the original reference paper.

Panel A: Financing

SV	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
CDI	0.13	0.11	(0.01, 0.05)	1.14	0.04	4.65	0.93	0.44	0.60
CSI	0.47	0.20	(0.00, 0.01)	0.92	0.57	2.99	1.00	0.99	0.52
DBE	0.39	0.20	(0.00, 0.06)	1.11	0.96	3.85	1.00	0.85	0.39
DCOL	0.16	0.17	(0.00, 0.10)	1.32	1.04	4.45	0.92	0.33	0.05
DFNL	0.32	0.16	(0.03, 0.14)	1.50	0.73	3.28	1.00	0.96	0.51
NDF	0.31	0.14	(0.06, 0.10)	1.38	0.69	3.39	1.00	0.99	0.80
NEF	0.33	0.20	(0.01, 0.03)	1.04	0.78	3.70	0.99	0.55	0.39
NXF	0.46	0.22	(0.03, 0.06)	1.24	0.73	3.54	1.00	0.91	0.63
Mean	0.32	0.17	(0.02, 0.07)	1.21	0.69	3.73	0.98	0.75	0.49

Panel B: Intangibles

SV	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
ADM	0.27	0.12	(0.00, 0.00)	0.51	0.68	3.48	1.00	0.17	0.40
CFV	0.33	0.17	(0.00, 0.00)	0.67	0.69	3.46	1.00	0.30	0.51
EPRD*	0.78	0.28	(0.03, 0.02)	1.08	0.44	2.42	1.00	1.00	0.93
HR	0.28	0.18	(0.02, 0.08)	1.28	0.67	4.02	0.99	0.67	0.33
KZI*	-0.01	0.09	(0.00, 0.00)	0.54	-0.17	4.95	0.50	0.00	0.02
LFE*	-0.04	0.07	(0.00, 0.01)	0.82	0.99	6.24	0.22	0.00	0.00
OL	0.29	0.12	(0.00, 0.00)	0.66	0.94	3.81	1.00	0.49	0.61
RDM	0.32	0.12	(0.00, 0.00)	0.64	1.48	7.02	1.00	0.26	0.01
RER	0.16	0.06	(0.00, 0.01)	0.80	0.84	4.07	1.00	0.86	0.54
TAN*	0.16	0.09	(0.00, 0.01)	0.72	0.48	5.23	0.95	0.12	0.41
WW*	-0.06	0.22	(0.00, 0.01)	0.83	-0.37	4.27	0.35	0.00	0.03
Mean	0.23	0.14	(0.01, 0.01)	0.78	0.61	4.45	0.82	0.35	0.35

Panel C: Investment

SV	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
ACI	0.25	0.11	(0.03, 0.03)	0.98	0.45	2.91	1.00	0.95	0.68
AG	0.48	0.26	(0.05, 0.11)	1.49	0.93	4.05	1.00	0.94	0.66
DINV	0.42	0.23	(0.09, 0.16)	1.71	0.78	3.15	1.00	1.00	0.81
DNOA	0.56	0.23	(0.09, 0.13)	1.63	0.85	3.76	1.00	1.00	0.85
DPIA	0.48	0.24	(0.09, 0.12)	1.60	0.80	3.83	1.00	0.95	0.68
DWC	0.45	0.22	(0.15, 0.13)	1.72	0.57	3.01	1.00	1.00	0.66
IG	0.34	0.15	(0.03, 0.05)	1.19	0.78	3.83	1.00	0.96	0.76
NOA	0.47	0.17	(0.02, 0.03)	1.05	0.87	4.08	1.00	1.00	0.45
OA	0.37	0.19	(0.01, 0.13)	1.42	0.78	3.46	1.00	0.88	0.44
PTA	0.31	0.16	(0.04, 0.05)	1.15	0.27	2.50	1.00	0.92	0.73
Mean	0.41	0.19	(0.06, 0.09)	1.40	0.71	3.46	1.00	0.96	0.67

Continued on next page

Table 2: Non-standard errors across sorting variables.

Panel D: Momentum

SV	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
52W	0.46	0.49	(0.16, 0.03)	1.44	-0.31	2.65	0.90	0.63	0.51
ABR	0.59	0.62	(0.21, 0.35)	4.12	0.98	2.73	0.99	0.92	0.73
MOM	0.57	0.45	(0.04, 0.07)	1.34	0.44	2.66	0.99	0.65	0.69
RMOM	0.43	0.26	(0.11, 0.03)	1.25	0.08	2.36	1.00	0.81	0.68
RS	0.39	0.22	(0.02, 0.15)	1.52	0.97	3.86	1.00	0.87	0.72
SUE	0.50	0.38	(0.16, 0.25)	2.67	1.09	3.92	1.00	0.87	0.90
TES	0.45	0.31	(0.09, 0.24)	2.30	1.01	3.29	1.00	0.95	0.71
Mean	0.48	0.39	(0.11, 0.16)	2.09	0.61	3.07	0.98	0.81	0.70

Panel E: Profitability

SV	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
ATO*	0.20	0.14	(0.00, 0.01)	0.78	0.55	3.10	1.00	0.25	0.25
BL*	-0.03	0.07	(0.00, 0.00)	0.38	0.17	3.45	0.25	0.00	0.00
CBOP	0.60	0.25	(0.05, 0.12)	1.57	0.77	3.71	1.00	0.97	0.99
CTO*	0.13	0.11	(0.00, 0.00)	0.61	0.81	4.10	0.96	0.03	0.06
GPA	0.33	0.20	(0.00, 0.04)	0.97	0.75	3.09	1.00	0.62	0.75
O	0.05	0.09	(0.00, 0.00)	0.54	0.22	3.38	0.79	0.02	0.01
OPE	0.33	0.14	(0.02, 0.00)	0.79	0.71	4.18	1.00	0.74	0.67
ROA	0.48	0.30	(0.05, 0.13)	1.52	0.99	3.79	1.00	0.83	0.73
ROE	0.50	0.27	(0.06, 0.15)	1.60	1.01	3.79	1.00	0.91	0.86
TBI	-0.09	0.27	(0.00, 0.03)	0.95	0.30	3.22	0.33	0.01	0.03
Z*	0.05	0.12	(0.00, 0.01)	0.63	1.33	5.97	0.65	0.01	0.00
Mean	0.23	0.18	(0.02, 0.05)	0.94	0.69	3.80	0.82	0.40	0.40

Panel F: Size

SV	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
ME	0.09	0.26	(0.11, 0.05)	1.50	1.02	9.51	0.70	0.08	0.17

Panel G: Trading frictions

SV	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
AMI	0.15	0.19	(0.00, 0.03)	1.04	0.75	8.15	0.85	0.18	0.25
BETA	0.01	0.09	(0.00, 0.00)	0.27	0.06	2.92	0.55	0.00	0.00
BFP	0.11	0.09	(0.00, 0.00)	0.26	0.13	3.69	0.95	0.00	0.00
DTV	0.30	0.17	(0.00, 0.02)	0.82	1.04	6.83	0.98	0.34	0.32
ISKEW	-0.02	0.12	(0.16, 0.05)	1.49	0.00	3.53	0.45	0.04	0.15
IVOL	0.30	0.22	(0.00, 0.03)	0.92	0.82	4.68	0.97	0.23	0.02
MDR	0.23	0.30	(0.00, 0.13)	1.16	1.05	3.75	0.93	0.24	0.07
SREV	0.03	0.99	(0.07, 0.33)	3.97	1.04	3.15	0.33	0.31	0.30
TUR	0.23	0.17	(0.00, 0.01)	0.70	0.69	4.71	0.94	0.09	0.08
Mean	0.15	0.26	(0.03, 0.07)	1.18	0.62	4.60	0.77	0.16	0.13

Panel H: Valuation

SV	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
AM*	0.24	0.17	(0.00, 0.01)	0.68	1.22	4.65	1.00	0.09	0.14
BM	0.33	0.19	(0.00, 0.03)	0.84	1.24	4.87	1.00	0.31	0.53
CFM	0.39	0.18	(0.00, 0.01)	0.69	0.69	3.21	1.00	0.57	0.85
DM	0.11	0.10	(0.00, 0.00)	0.42	0.57	3.34	0.96	0.00	0.06

Continued on next page

Table 2: Non-standard errors across sorting variables.

EBM	0.23	0.14	(0.00, 0.01)	0.63	1.28	5.64	0.99	0.09	0.52
EM	0.36	0.11	(0.00, 0.00)	0.57	0.87	4.08	1.00	0.71	0.75
NDM	0.07	0.08	(0.00, 0.00)	0.44	0.67	4.19	0.90	0.00	0.11
NPY	0.26	0.07	(0.00, 0.00)	0.44	0.47	3.51	1.00	0.72	0.77
OCM	0.43	0.14	(0.00, 0.00)	0.59	0.48	3.13	1.00	0.79	0.72
REV	0.19	0.13	(0.00, 0.02)	0.77	1.45	7.30	0.98	0.14	0.43
SM	0.42	0.20	(0.00, 0.01)	0.78	1.14	4.17	1.00	0.52	0.90
Mean	0.28	0.14	(0.00, 0.01)	0.62	0.91	4.37	0.98	0.36	0.52

Panel I: Overall

SV	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
All	0.29	0.20	(0.03, 0.06)	1.12	0.70	4.05	0.90	0.52	0.45
Orig. Sig.	0.31	0.21	(0.04, 0.07)	1.19	0.73	3.99	0.94	0.58	0.50
Orig. Insig.	0.14	0.13	(0.00, 0.01)	0.71	0.54	4.44	0.69	0.15	0.18

3.2 Factor-adjusted premia

It is a standard procedure in asset pricing to control for factor exposures of various models. Thereby, we can infer how much of the measured premia are actually due to exposure to other, well-established risk factors. In particular, some return differentials can be fully explained by other risk factors, i.e., they represent an indistinguishable return pattern. In this section, we investigate non-standard errors of premia adjusted for their exposure to the CAPM, the Fama and French (2015) five-factor model (abbreviated FF5), and the Hou et al. (2021) five-factor model (abbreviated Q5). This question is relevant as part of the variation induced by varying over all decision nodes could be systematically related to common factor returns. Thus, the variation in the factor exposures across specifications could give rise to non-standard errors that could be alleviated by controlling for standard factors.

Figure 5 shows that non-standard errors in CAPM alphas are, on average, sizeable and comparable to their unadjusted counterparts shown in Figure 4. For the majority of sorting variables, the variation in CAPM alphas even exceeds the variation observed in the unadjusted premia. The average interquartile range across the distributions of all sorting variables is 0.22% per month, as shown in Table 3. This compares to 0.20% for unadjusted premia. The elevated variation is accompanied by an increase in the mean premium to 0.38% per month for CAPM alphas in Table II.1. Controlling for the exposure to market risk also increases the average frequency of statistically significant specifications for predictors that significantly relate to mean returns in the original paper from 58% for unadjusted premia to 74% for CAPM alphas in Table II.1. Since non-standard errors are, on average, larger for CAPM alphas but standard errors are lower for CAPM-adjusted returns, the ratio of the standard deviation of premia to average standard errors is larger for CAPM alphas compared to unadjusted premia.

We also investigate non-standard errors in alphas from the two five-factor models of Fama and French (2015) and the Hou et al. (2021) as mentioned above. Table 3 reports the non-standard errors for factor-adjusted alphas. First, we report average non-standard errors within sorting variable groups for the unadjusted premia and the factor model alphas. Then, we show average Anderson-Darling test statistics to assess whether the resulting distributions of the (un-)adjusted premia are different from each other.

Table 3 shows that the previously discussed increase in non-standard errors for CAPM alphas is observable for all groups except for size. Moreover, the Anderson-Darling test statistic in column 'R-C' shows that the distributions of unadjusted ("raw") premia and CAPM alphas are significantly different from each other.[11] The non-standard errors of FF5 alphas are, on average, and across all groups quantitatively similar to the non-standard errors of unadjusted premia in Table 3. This suggests that non-standard errors are not related to FF5 factors and is surprising because controlling for FF5 factor exposure reduces the level and significance of premia for nearly all groups in Table II.2 in the Internet Appendix. Only controlling for Q5 factor exposure reduces the overall non-standard error slightly by 0.02 percentage points. This reduction relative to unadjusted premia is particularly pronounced for sorting variables belonging to the groups of momentum, investment, financing, and profitability. Finding that non-standard errors are lower for momentum-related predictors is particularly remarkable because the Q5 model does not include a momentum factor. In contrast, sorting variables in the groups of valuation and trading frictions show hardly any decrease in non-standard errors. Relative to unadjusted premia, size or predictors associated with the group of intangibles have, on average, even higher non-standard errors when controlling for the Q5 model.[12]

[11] In fact, all Anderson-Darling test statistics in Table 3 correspond at least to the 1% significance level evaluated relative to a studentized distribution.

[12] We provide the detailed non-standard errors and summary statistics for premia relative to the Hou et al. (2021) model in Table II.3 in Internet Appendix II.

Figure 5: CAPM alphas: Non-standard errors across sorting variables.
This figure shows the estimated average premia (in %) adjusted for the CAPM in box plots for all sorting variables across all decision nodes. The vertical axis shows the associated sorting variable, while the color scheme connects each sorting variable to the respective category.

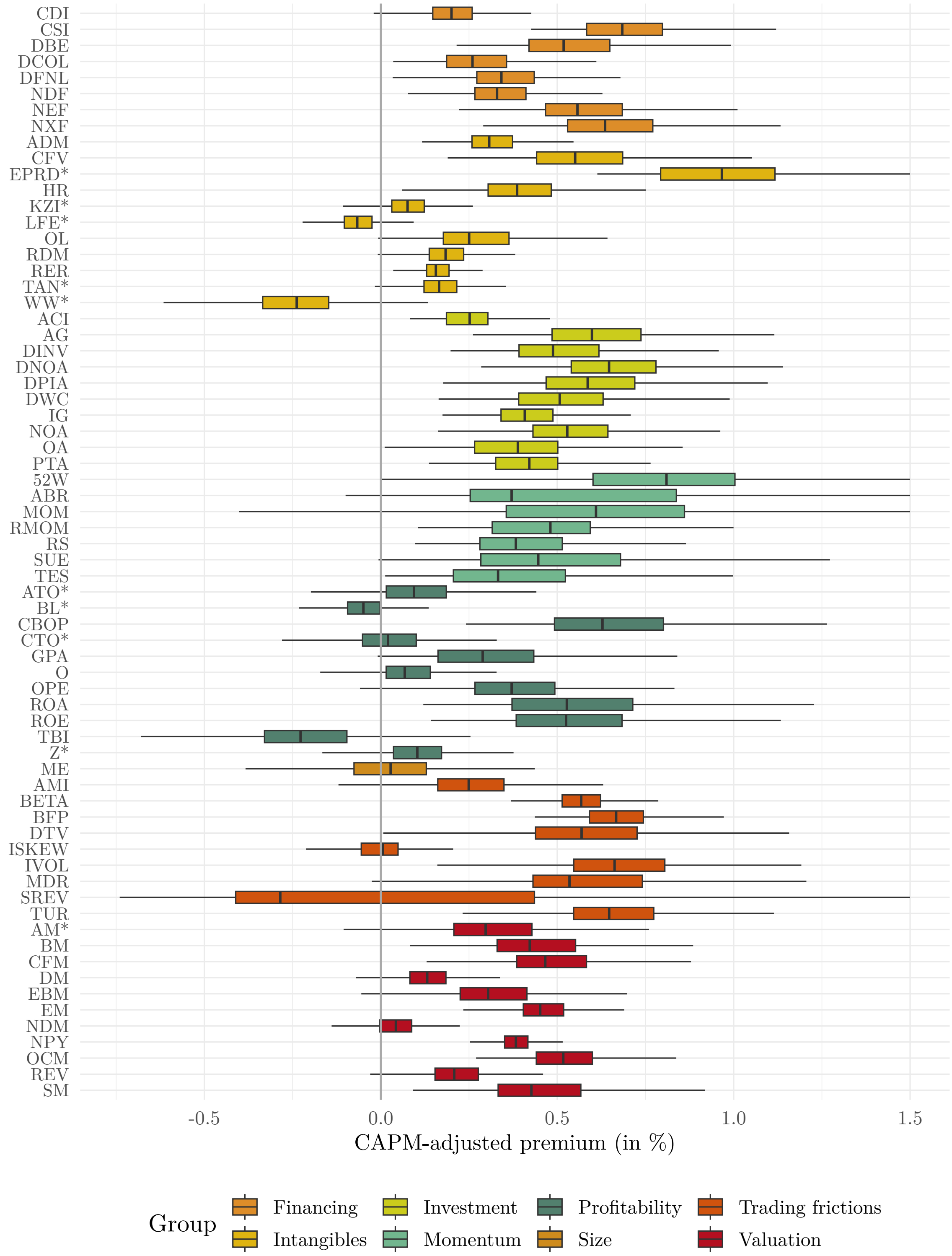


Table 3: Non-standard errors across asset-pricing models.
This table shows non-standard errors for average premia (Raw, in %), CAPM-adjusted premia (CAPM, in %), FF5-adjusted premia (FF5, in %), Q5-adjusted premia (Q5, in %). Then, we test the similarity of the demeaned distributions of the models using Anderson-Darling tests (Scholz and Stephens, 1987). We report the average test statistics between unadjusted ("raw") and CAPM-adjusted premia (R-C), premia of the CAPM against the FF5 (C-F) and Q5 model (C-Q), and between the FF5 and Q5 model (F-Q).

Group	Raw	CAPM	FF5	Q5	R-C	C-F	C-Q	F-Q
Financing	0.17	0.19	0.17	0.14	97	433	719	408
Intangibles	0.14	0.15	0.17	0.15	320	1137	718	343
Investment	0.19	0.21	0.19	0.15	164	442	1224	779
Momentum	0.39	0.40	0.38	0.31	86	150	936	679
Profitability	0.18	0.21	0.16	0.15	669	1269	1355	374
Size	0.26	0.21	0.18	0.42	142	90	1203	1417
Trading frictions	0.26	0.28	0.25	0.25	264	122	266	128
Valuation	0.14	0.16	0.14	0.13	286	869	1034	182
Mean	0.20	0.22	0.20	0.18	287	679	917	416

Next, we consider differences among factor models in addressing non-standard errors. Indeed, we find that both the FF5- and Q5-adjusted premia have lower average non-standard errors than under the CAPM. This reduction of non-standard errors is stronger for the Q5 model compared to the FF5 model. Thus, the distributions of Q5 alphas and CAPM alphas differ more in Table 3 relative to the comparison of FF5 and CAPM alphas. Apart from differences in non-standard errors, we also find large differences in the average alphas and in the share of significant factor alphas. Compared to an average unadjusted premium of 0.29% per month, average FF5 alphas and Q5 alphas are 0.21% and 0.09% per month (Table II.2 and Table II.3 in the Internet Appendix). However, this reduction is primarily due to variables upon which the factors are based. Overall, we find that the Q5 model picks up more variation in estimated premia and yields fewer significant alphas (only 21% of specifications report a t-value larger than 1.96) compared to the FF5 model (50% significant).

In a nutshell, non-standard errors are also considerably large and statistically significant when controlling for established factor models. This suggests that controlling for factor exposure does by no means alleviate non-standard errors. While the methodology-induced variation in CAPM alphas is even larger compared to "raw" premia, it is unchanged for Fama and French (2015) factor alphas. Only alphas relative to the Hou et al. (2021) model show a slight reduction in non-standard errors for some groups of sorting variables.

4 Which methodological choices induce the largest variation in estimated premia?

The large non-standard errors in portfolio sorts highlighted in the previous section motivate us to investigate the impact of the 14 methodological decision nodes directly. We study the nodes' impacts on two levels. First, we analyze which decision nodes generally induce the largest variation in premia when all decisions but the one in the respective node are kept constant, i.e., we analyze the impact *across branches*. For instance, we answer the question by how much premia differ solely depending on whether or not one excludes small stocks but when all variation generated by other decision nodes is taken out. Second, we study the variation in estimated premia when a specific choice in a decision node is kept constant, i.e., we measure non-standard errors *within branches*. For example, we study how much variation in premia there can be if we exclude small caps but do not fix whether or not one uses NYSE breakpoints, includes firms with negative book equity or not, and so on. Finally, we investigate the same questions for CAPM-adjusted returns.

4.1 Impact across branches

We start by comparing premia that only differ in one decision node. We focus on two main outcome variables, the mean absolute difference in the time series of premia and the average correlation between premium estimates. Panel A of Table 4 shows the mean absolute differences across branches. These are calculated as the time-series average of the quantity in Equation (5), i.e., the mean of pairwise absolute differences between premia with specifications that are identical except for one decision node (which is specified in the first column of Panel A). For instance, the first row shows the difference between the premia in all specifications that are computed using identical decisions in all nodes except for the choice of quantile breakpoints (i.e., quintiles or deciles). Intuitively, if the mean absolute difference in Panel A of Table 4 is high, the node under investigation induces large variation in estimated premia even if all other decisions are kept constant.

Overall, we find large impacts in terms of mean absolute differences, which in most cases exceed half a percentage point per month. A few cases stand out: As expected, the number of portfolios has a large impact on non-standard errors. The weighting scheme is similarly important as conditioning on positive earnings. Moreover, the size restriction, the sorting variable lag, and the exchanges used for setting quantile breakpoints (particularly so for sorts on size) induce considerable and quantitatively similar variation in estimated premia. This shows that other nodes besides the size-related nodes discussed in Hou et al. (2020) induce large variation in premia.

The mean absolute differences vary across groups of sorting variables indicated by the remaining columns in Panel A. Profitability anomalies have large mean absolute differences depending on whether only firms with positive earnings are included, most likely because profitability is related to earnings. Similarly, the exclusion of financial firms has a strong effect on profitability-related return differentials. Moreover, the sorting variable lag tends to have a particularly strong impact for sorting variables belonging to momentum or trading frictions, where cross-sectional predictability is rather ephemeral. While the exclusion of utilities does not have a strong impact on sorts based on intangibles, investment, financing, and momentum, its effect amounts to more than half a percentage point for profitability, size, trading friction, and valuation sorts. In short, there is large heterogeneity in the impact of decision nodes for different groups of sorting variables. This suggests that merely restricting size-related decision nodes is not enough.

A different way to consider the impact of decision nodes is to compare the time-series correlation of estimated premia. Panel B of Table 4 shows the correlation of the time series of return differentials. As in Panel A, the decisions in all nodes but the node specified in the first column are kept constant. For instance, the second row in Panel B shows the average pairwise correlation between return time series that only differ in whether they include or exclude firms with negative earnings (see Equation (6)). A lower correlation between return differentials indicates a stronger impact of the respective decision node. This also has an economically meaningful interpretation. If a long-short strategy captures robust exposure to a (risk) factor, small alterations in the trading strategy should not affect the factor structure, and hence, all such strategies should have highly correlated returns. Thus, stable correlations across branches are indicative of whether an anomaly represents a factor. The ranking of decision nodes in terms of impact is similar to the one based on mean absolute differences in Panel A. However, the decision to exclude negative earnings stocks and the choice of the sorting variable lag are now among the three most impactful decision nodes. Comparing across groups of sorting variables, momentum and valuation stand out in their non-susceptibility with respect to changes in decision nodes when measured in terms of correlations, which are high across branches. This indicates that irrespective of the decision nodes, momentum and valuation sorts exhibit a great deal of common variation, even when the mean absolute differences are large, as shown in Panel A. This points to momentum and valuation having a strong factor structure robust to the exact choices of sample composition and portfolio formation. At the opposite end of the spectrum in terms of stability are sorting variables belonging to investment, financing, and size, which are particularly susceptible to losing correlation when altering choices about rebalancing or excluding stocks with negative earnings.

Lastly, we test whether the distributions of premia which only differ in the node under investigation, are drawn from the same distribution. Therefore, we implement an Anderson-Darling test following Scholz and Stephens (1987) and report the corresponding test statistic in Panel C of Table 4.

The larger the test statistic in Panel C, the more dissimilar the distributions of premia for the decision node under investigation. Thus, larger test statistics imply a larger impact on estimated premia relative to other decision nodes. The results in Panel C are similar to mean absolute differences and correlations of premia. Even more pronounced than for size-related decision nodes, the choices for the sorting variable lag and the exclusion of stocks with negative earnings create distributions of premia that differ substantially. Similar to correlations of premia in Panel B, we find that the test statistics are considerably low for sorting variables belonging to momentum, indicating stability with respect to methodological choices. Contrarily, choices for most of the decision nodes induce large differences in the distributions for predictors belonging to financing and investment.

In a nutshell, size-related decision nodes discussed in Hou et al. (2020) induce considerable variation in estimated premia. However, we find that the inclusion of stocks with negative earnings and the sorting variable lag tend to induce even larger variation. Moreover, we observe considerable heterogeneity in the impact of decision nodes on premia from different groups. Therefore, we encourage researchers to take this heterogeneity into account when analyzing premia from a specific group.

4.2 Effects within branches

We now investigate how much variation in premia is removed when taking a specific decision in one node but allowing for variation in all other nodes. In particular, we consider the distribution of premia within branches and thereby assess the *impact of specific choices* in a decision node rather than the *impact of the node* when controlling for all other decisions. This analysis reveals which exact choice of a decision node induces larger non-standard errors in the remaining decision nodes. Analyzing the variation conditional on a specific choice of a decision node allows us to study the robustness of our results to a specific decision node. For example, it could be that once we decide to exclude penny stocks, it does not matter much what other decisions we take. We present this within-branch variation in Table 5 in terms of mean premia and non-standard errors similar to the statistics reported in Table 2. Moreover, we focus on the nodes with the largest impact in terms of mean absolute differences in premia in Table 4 and show the remaining nodes in the Internet Appendix.

Table 4: Mean absolute differences and correlations across decision nodes.
This table shows mean absolute differences (Panel A, in %) and correlations (Panel B) of the time series of premia across individual decision nodes. For each decision node, we compare time-series pairs that differ only in the specific node. Then, we take the mean for each node-sorting variable combination. In Panel C, we show average Anderson-Darling test (Scholz and Stephens, 1987) comparing the distribution of premia differing only in one decision node. The three panels show means for all categories together (Overall) and individual categories separately. Moreover, the nodes are arranged by impact. By construction, some entries do not produce variation and are left empty.

Panel A: Mean absolute differences

Node	All	Fin.	Int.	Inv.	Mom.	Pro.	Siz.	Tra.	Val.
BP: Quantiles (main)	1.08	0.90	1.12	0.96	1.04	1.08	1.40	1.37	1.02
Weighting scheme	1.00	0.96	1.05	0.96	1.01	1.02	0.67	0.92	1.08
Positive earnings	0.98	0.86	0.99	0.85	0.79	1.28	1.37	1.07	0.87
Size restriction	0.86	0.68	0.85	0.69	0.79	0.90	1.95	1.15	0.80
Sorting variable lag	0.84	0.56	0.53	0.63	1.69	0.64	1.85	1.53	0.57
BP: Exchanges	0.84	0.70	0.87	0.68	0.67	0.90	1.60	1.15	0.79
Financials	0.75	0.46	0.75	0.61	0.60	1.06	0.81	0.76	0.86
Double sort	0.70	0.42	0.67	0.45	0.51	0.74	2.97	1.33	0.52
BP: Quantiles (second)	0.70	0.53	0.69	0.53	0.59	0.71	1.86	1.07	0.61
Rebalancing	0.58	0.59	0.58	0.62		0.59			0.54
Utilities	0.50	0.37	0.37	0.33	0.42	0.72	0.60	0.66	0.57
Stock-age restriction	0.42	0.45	0.39	0.43	0.32	0.47	0.64	0.52	0.35
Price restriction	0.37	0.31	0.39	0.32	0.32	0.40	0.77	0.48	0.35
Positive book equity	0.07	0.07	0.08	0.07	0.07	0.08	0.13	0.10	0.06

Panel B: Correlations

Node	All	Fin.	Int.	Inv.	Mom.	Pro.	Siz.	Tra.	Val.
Weighting scheme	0.87	0.82	0.85	0.82	0.88	0.88	0.97	0.93	0.90
Positive earnings	0.88	0.86	0.88	0.83	0.93	0.80	0.84	0.93	0.93
Sorting variable lag	0.89	0.92	0.95	0.91	0.72	0.95	0.63	0.76	0.96
BP: Quantiles (main)	0.91	0.88	0.88	0.87	0.92	0.92	0.88	0.92	0.94
Size restriction	0.92	0.91	0.91	0.90	0.93	0.92	0.74	0.91	0.95
Financials	0.92	0.96	0.90	0.91	0.96	0.85	0.93	0.96	0.94
BP: Exchanges	0.93	0.92	0.92	0.92	0.95	0.93	0.81	0.92	0.96
Rebalancing	0.93	0.90	0.93	0.90		0.95			0.96
BP: Quantiles (second)	0.94	0.94	0.94	0.94	0.96	0.95	0.72	0.92	0.97
Double sort	0.94	0.96	0.95	0.96	0.97	0.94	0.63	0.89	0.98
Utilities	0.97	0.97	0.98	0.98	0.98	0.93	0.95	0.98	0.97
Stock-age restriction	0.97	0.95	0.97	0.96	0.99	0.97	0.94	0.98	0.99
Price restriction	0.97	0.97	0.97	0.97	0.98	0.97	0.93	0.98	0.98
Positive book equity	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00

Panel C: Anderson-Darling test statistics

Node	All	Fin.	Int.	Inv.	Mom.	Pro.	Siz.	Tra.	Val.
BP: Quantiles (main)	3180	3414	3714	8201	1053	2391	232	629	2407
Positive earnings	3105	2310	3515	2627	399	6171	136	390	4853
Sorting variable lag	3068	2743	440	4685	9278	1560	4027	4357	877
Size restriction	2441	3157	3031	2351	958	1887	1372	927	4245
Weighting scheme	2331	5281	3180	2233	354	1079	48	886	3323
BP: Exchanges	2120	3746	1341	2868	522	1859	213	826	3548
Rebalancing	1679	2768	1059	3151		1335			375
Double sort	1380	1052	2204	330	490	2227	3630	1501	1162
Financials	723	276	1266	795	96	1145	6	177	929
Stock-age restriction	672	745	974	1156	139	781	72	358	342
Utilities	369	126	75	182	85	930	17	528	531
BP: Quantiles (second)	306	221	595	63	40	434	1271	123	402
Price restriction	218	144	300	158	40	205	424	131	422
Positive book equity	7	4	10	5	0	11	0	1	13

Impact of: Breakpoint quantiles (main). Panel A of Table 5 shows the non-standard errors for branches of the most impactful decision node from above. Across all sorting variables, non-standard errors and the ratios of the dispersion of estimated premia to average standard errors are higher for decile breakpoints (Panel A). Moreover, a larger amount of premia are monotonic for quintiles compared to decile portfolios. These findings are similar for most groups of sorting variables in Table C.1, if we average across sorting variables belonging to the respective group. If a sorting variable is monotonically related to subsequent mean returns, return differentials should be higher for more extreme quantiles. Consequently, small methodological changes in other nodes may induce larger non-standard errors in decile portfolios as opposed to quintile portfolios.

Impact of: Weighting scheme. Next, we turn to the impact of different weighting schemes, i.e., whether portfolio returns are value or equally weighted. Results are presented in Panel B of Table 5. Mean premia are higher for equally weighted returns, as are non-standard errors. This is probably the case because smaller firms have more exposure to a variety of priced risks (and potential mispricing). These stocks get a smaller weight in value-weighted sorts. Although there are large differences when considering all sorting variables jointly, non-standard errors between value and equally weighted returns in Table C.2 are almost identical for sorting variables belonging to intangibles, investment, and profitability. This points to a factor structure for these groups being equally strong for small and large stocks within portfolios.

Impact of: Positive earnings filter. Panel C of Table 5 shows the results for the decision node that considers restricting the sample to firms with positive earnings. Across all sorting variables, mean premia as well as non-standard errors are considerably lower if we exclude firms with negative earnings. Moreover, there is large heterogeneity across groups: There are no notable differences for sorting variables belonging to momentum or trading frictions. Contrary, we find that smaller non-standard errors for samples excluding firms with negative earnings are particularly pronounced for sorting variables belonging to financing, intangibles, or profitability in Table C.3. The effect on profitability premia is intuitive because stocks with negative earnings have very low profitability.

Table 5: Impact of specific choices on premia.
This table shows summary statistics holding the individual choices of the panel's decision node constant. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. We also show the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: BP: Quantiles (main)

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
5	0.25	0.17	(0.02, 0.06)	1.03	0.66	3.75	0.90	0.51	0.45
10	0.32	0.22	(0.02, 0.06)	1.07	0.63	3.82	0.90	0.52	0.33

Panel B: Weighting scheme

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
EW	0.31	0.21	(0.04, 0.06)	1.16	0.73	3.90	0.90	0.57	0.47
VW	0.26	0.18	(0.02, 0.05)	1.01	0.67	3.94	0.90	0.46	0.44

Panel C: Positive earnings

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.31	0.21	(0.03, 0.05)	1.09	0.66	3.90	0.91	0.53	0.44
Yes	0.26	0.18	(0.03, 0.06)	1.05	0.53	3.50	0.89	0.50	0.47

Panel D: Size restriction

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.33	0.24	(0.05, 0.08)	1.28	0.60	3.54	0.90	0.57	0.48
0.2	0.24	0.16	(0.01, 0.03)	0.86	0.35	3.17	0.89	0.45	0.44

Panel E: Sorting variable lag

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
1m	0.28	0.17	(0.02, 0.04)	0.95	0.63	5.09	0.88	0.34	0.25
3m	0.28	0.16	(0.03, 0.04)	0.95	0.63	4.15	0.89	0.49	0.44
6m	0.25	0.15	(0.01, 0.03)	0.86	0.59	4.06	0.88	0.48	0.42
FF	0.26	0.14	(0.01, 0.03)	0.86	0.68	3.80	0.91	0.52	0.45

Panel F: BP: Exchanges

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
All	0.32	0.22	(0.04, 0.06)	1.17	0.56	3.60	0.90	0.54	0.46
NYSE	0.26	0.17	(0.02, 0.04)	0.96	0.63	4.06	0.90	0.49	0.45

Impact of: Size restriction. Mean return differentials are smaller when excluding the smallest 20% of stocks at each formation time, as shown in Panel D of Table 5. Moreover, non-standard errors are lower, as are the ratios of non-standard errors to average standard errors. This suggests that many sorting variables are related to market capitalization. Moreover, the share of sorts that generate monotonically increasing mean returns also decreases when we exclude smaller stocks, similar to value weighting portfolios in Panel B. The reduction in monotonicity is particularly high for sorting variables belonging to the groups of intangibles, trading frictions, and valuation, as documented in Table C.4.

Impact of: Sorting variable lag. Another important decision node unrelated to size is the selection of the time lag between the arrival of information about the sorting variable and portfolio formation, which we analyze in Panel E of Table 5. This decision node is not binary because we allow for lags of one, three, six, or at least six months as in Fama and French (1992), depending on the update frequency of the sorting variable. Thus, not all four choices are available for all sorting variables.[13] Mean premia and non-standard errors are larger for shorter time lags regardless of the sorting variable's update frequency. This indicates that some cross-sectional predictability can only be exploited over short horizons because information is incorporated into prices rather quickly. Interestingly, the choice of a specific sorting variable lag also has important implications for the monotonicity of premia as documented in Table C.5. Although shorter lags induce more uncertainty about the size and significance of the premia, they also lead to more monotonic relations between the predictor and subsequent mean returns for some groups. In summary, we find the strongest impact in terms of mean premia, non-standard errors, and monotonicity for sorting variables belonging to the groups investment, momentum, profitability, and trading frictions. Hence, we show that the sorting variable lag is not only important for the book-to-market ratio as in Asness and Frazzini (2013) but particularly for sorting variables belonging to investment, momentum, profitability, and trading frictions.

Impact of: Breakpoint exchanges. We investigate the decision to base breakpoints on specific exchanges such as the NYSE in Panel F of Table 5. The results are similar to the size-related decision nodes "weighting scheme" and "size restriction". Notably, we find that mean premia, non-standard errors, and the ratio of the dispersion of premia to the average time-series standard error are smaller if we calculate breakpoints only with stocks listed on NYSE. For this particular choice, the fraction of monotonic premia is notably reduced for sorting variables belonging to trading frictions and valuations in Table C.6.

[13] We investigate a lag of one, three, or six months for predictors which are updated monthly, a lag of three or six months for sorting variables updated quarterly, and a lag of three, six, or at least six months (Fama and French, 1992) for sorting variables updated on an annual basis.

Impact of: Other nodes. Finally, we consider decision nodes with a mean absolute difference across branches below 0.8 percentage points in Table 4. That said, the effects on other diagnostics may be substantial. For instance, while the choice to exclude financial stocks has small effects on mean premia and non-standard errors, it has large and heterogeneous effects on the monotonicity of premia in Table C.7: Return premia from sorting variables related to financing, intangibles, and valuation tend to be less monotonic if stocks in the financial sector are excluded. Contrarily, investment premia are more often monotonic if financial stocks are excluded. Therefore, although the choice to exclude stocks from the financial sector has only a small impact on non-standard errors, it has a profound impact on the monotonicity of mean returns. Similar to size-related decision nodes, the choice to implement a double sort with size as the secondary sorting variable reduces mean premia and non-standard errors in the remaining decision nodes in Table C.8 for almost all groups. The choice to rebalance portfolios monthly compared to a yearly rebalancing frequency introduces larger mean premia and non-standard errors for almost all groups of sorting variables in Table C.10. Even though the uncertainty about the sign and significance of premia is larger for monthly rebalancing, the monotonicity of premia is more pronounced. These findings are similar to the decision node "sorting variable lag" because also the rebalancing frequency decides how "fresh" the sorting variable information is at the time of portfolio formation.[14] The exclusion of utility stocks leads to almost identical mean premia and non-standard errors in the remaining decision nodes, as compared to specifications including utility stocks. However, similar to the decision node "financials", the impact is more nuanced as the monotonicity of return premia varies widely along this choice in Table C.11. For instance, sorts based on financing, investment, profitability, and size variables are less often monotonic if utility stocks are excluded. The reverse holds for sorting variables belonging to the groups of trading frictions and valuation. On the other hand, we observe only negligible differences for non-standard errors in the choices belonging to the decision nodes: "breakpoint quantiles (secondary)", "stock age restriction", "price restriction", and "positive book equity" in the corresponding Tables C.9 through C.14.

Overall, these results have the following three implications: First, our results of large non-standard errors discussed in the previous section are robust to keeping specific nodes constant. For instance, even if we hold any of the six most decisive nodes in Table 5 constant, we still observe statistically significant non-standard errors above 0.15% per month. Second, decision nodes such as "financials" and "utilities" with negligible impact on non-standard errors of premia in the previous section reveal a strong impact on the monotonicity of mean portfolio returns. While this might be less important from

[14] Unlike for the "sorting variable lag", the interpretation is less clear cut since the effect is mostly driven by differences in fiscal year-end months for firms in our sample.

a mispricing perspective, it is relevant if the relation to expected returns is theoretically motivated by factor exposures. Third, although decisions to reduce the impact of small stocks lead to smaller non-standard errors in the remaining decision nodes, they also reduce the monotonicity of premia. Therefore, one loses interesting and potentially risk-related variation if size-related decision nodes are fixed in favor of stocks with large market values.

4.3 The impact of individual decision nodes on alphas

In this section, we investigate which nodes drive the variation in adjusted returns in the time series and how large non-standard errors in alphas are in the respective node's branches. To analyze the effect of each node on the time series of adjusted returns, we run time-series regressions of the monthly premium for each sorting variable v and each specification s on the respective factor model $M \in \{\text{CAPM}, \text{FF5}, \text{Q5}\}$. Thereby, we obtain the intercept $\tilde{\alpha}_s^{M,v}$ and the residuals $\tilde{\varepsilon}_{t,s}^{M,v}$. In a second step, we add the respective intercept and the residuals of these time-series regressions, i.e., we define the *adjusted return* time series as:

$$\tilde{ar}_{t,s}^{M,v} := \tilde{\alpha}_s^{M,v} + \tilde{\varepsilon}_{t,s}^{M,v}. \tag{7}$$

The time-series variation is then driven by the residuals and the level difference in alphas, respectively. We investigate the sum of alphas and residuals from regressing premia on factor models because we are interested in *all* the variation that is unexplained by factor models.

Similar to Table 4, we show the effects of individual decision nodes on the time series of CAPM alphas in Table D.1.[15] Exactly as in Section 4 above, we only compare time series that differ in exactly one decision node. The mean absolute differences in Panel A of Table D.1 and the correlations in Panel B for CAPM alphas are almost identical compared to unadjusted premia in Table 4. Additionally, the order of the most decisive decision nodes is unchanged. The same holds when we analyze FF5- and Q5-adjusted returns in Tables D.2 and D.3. This indicates that differences in premia for specific choices of a decision node are not systematically related to factors proposed by the CAPM, FF5 model, or the Q5 factor model.

Finally, we present an assessment of the variation in adjusted returns (i.e., non-standard errors) for the branches of the decision nodes with the highest mean absolute differences. Table VI.2 shows the non-standard errors averaged across all sorting variables (irrespective of their group) for CAPM alphas.[16] The results for CAPM alphas are in line with results for unadjusted premia presented in

[15] We show results for the FF5-adjusted and Q5-adjusted returns in the Appendix in Tables D.2 and D.3. Additional summary statistics for both models are in Internet Appendix Section VI.

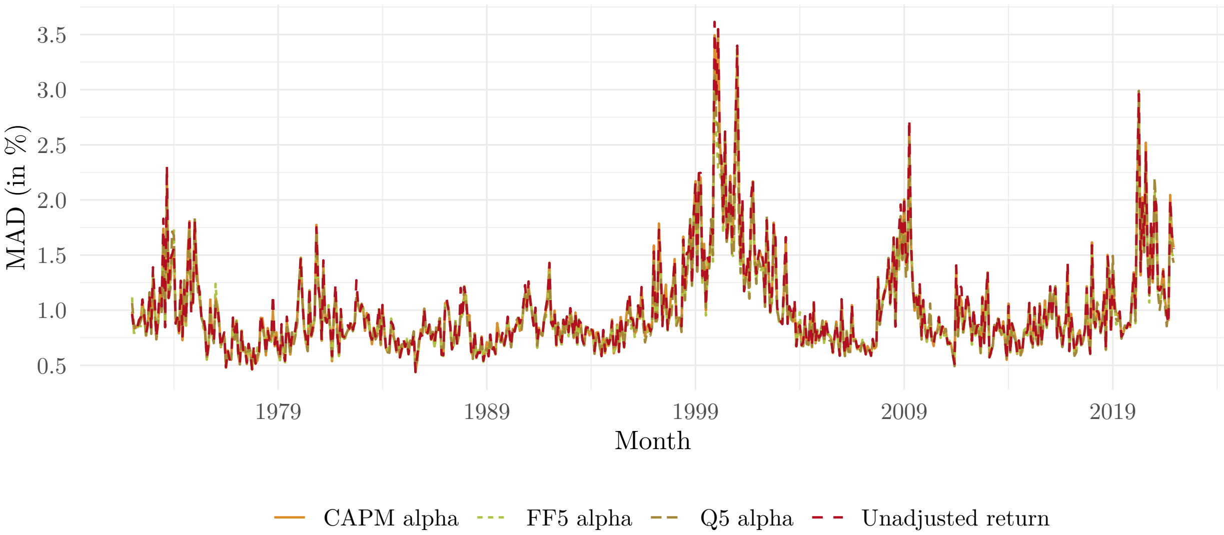
[16] We show the remaining decision nodes' impact on CAPM alphas in Table VI.3 in the Internet Appendix.

Section 4.2. Not surprisingly, the level of mean premia differs within branches compared to unadjusted premia. However, we observe very similar non-standard errors as well as fractions of monotonic premia for CAPM alphas in the respective branch compared to unadjusted premia. Again, the same holds if we evaluate FF5- and Q5-adjusted returns in Table VI.4 to Table VI.7. Overall, the main insights from the decision nodes' impact on premia are not changed when adjusting premia for their exposure to factor models.

5 Economic drivers of non-standard errors in portfolio sorts

Return differentials from portfolio sorts have significant non-standard errors. It is natural to ask if this variation in estimated premia has an underlying economic driver. As a starting point to answering this question, we investigate the variation over time of mean absolute differences from Equation (5). As an example, Figure 6 depicts the time series of mean absolute differences for the decision node "weighting scheme".

Figure 6: Mean absolute differences over time for the decision node "weighting scheme". This figure shows the time series of mean absolute differences (in %) for the decision node "weighting scheme". We plot the differences for unadjusted premia, CAPM, FF5, and Q5 alphas.



The mean absolute differences in Figure 6 for the unadjusted premia and several factor-adjusted returns exhibit very similar time series patterns. There are considerable increases in mean absolute differences around the oil crises (1973, 1979), the "dot-com bubble" (1999 - 2001), the financial crisis (2007 - 2009), and the Covid crisis (2019 - 2021). Note that we find similar time variation for most of the remaining 13 nodes in Figure VII.1 - Figure VII.13 in the Internet Appendix VII. All these periods tend to be associated with high market volatility. Intuitively, when market volatility is high, the cross-sectional dispersion of returns increases as well. Unsurprisingly, cross-sectional return differentials are heavily influenced by this.

To understand how non-standard errors vary over time above and beyond what is implied by dispersion, we regress the mean absolute differences for each decision node n on the cross-sectional return dispersion in each month:[17]

$$MAD_t^n = \alpha^n + \beta^n \cdot \sigma_t^{ret} + \text{Residual MAD}_t^n. \tag{8}$$

Thereafter, we regress the residuals from (8) on various economic indicators, such as the CBOE volatility index (VIX), the NBER recession indicator, the Pástor and Stambaugh (2003) liquidity index, and the Baker and Wurgler (2006) sentiment index.[18]

$$\text{Residual MAD}_t^n = \alpha^n + \beta^n \cdot \text{Economic indicator}_t + \epsilon_t^n \tag{9}$$

The regression coefficients of these time-series regressions in Table 6 confirm the patterns observed in Figure 6. Residual mean absolute differences for most nodes tend to be larger when markets are volatile, in a recession, or illiquid. For example, a one-standard-deviation change in the CBOE volatility index corresponds to a 0.40 standard deviation increase in mean absolute differences for the node "weighting scheme". Some nodes, such as "stock age" or "price restrictions", show almost no relation to these economic indicators. However, these are also the decision nodes with the smallest mean absolute differences over the full sample in Table 4. While the coefficients for the sentiment index from Baker and Wurgler (2006) are insignificant for most decision nodes, we find a statistically significant and positive coefficient for the decision nodes "positive earnings" and "breakpoint exchanges". Intuitively, stocks with positive rather than negative earnings or stocks listed on NYSE and AMEX compared to technology stocks on NASDAQ tend to have different exposures to sentiment. These cross-sectional differences in sentiment have also been associated with differences in mean returns (Baker and Wurgler, 2006). Therefore, the choices for these decision nodes potentially create large differences in subsequent mean returns as captured by our regression coefficients in Table 6.[19]

[17] We measure the cross-sectional return dispersion as the cross-sectional standard deviation of stock returns for all common U.S. stocks listed on NYSE, AMEX, and NASDAQ in a given month. We adjust returns for delisting according to Shumway (1997).

[18] Note that the CBOE volatility index, the NBER recession indicator, the liquidity index from Pástor and Stambaugh (2003), and the Residual MAD_t^n are stationary for most nodes n. For the remaining decision nodes, as well as for the regression with the sentiment index, we find a cointegration relation between the mean absolute differences and the economic indicators. An augmented Dickey-Fuller test (Dickey and Fuller, 1979) confirms that the residuals from (9) are stationary.

[19] Note that the regression coefficients are quantitatively and also statistically very similar if mean absolute differences are adjusted for factor models (untabulated).

Table 6: Time-series regressions of mean absolute differences on economic state variables. Each entry corresponds to a time-series regression of the residual mean absolute differences for the corresponding node on one of the following economic state variables: the CBOE volatility index, the NBER recession indicator, the liquidity index from Pástor and Stambaugh (2003), and the sentiment index from Baker and Wurgler (2006). We calculate residual mean absolute differences relative to the cross-sectional return dispersion in each month as in equation (8). Moreover, mean absolute differences are calculated as in equation (5) and are based on unadjusted premia. Cross-sectional return dispersion corresponds to the standard deviation of returns of U.S. stocks from CRSP in each cross-section. The dependent and independent variables are standardized, Newey and West (1987) corrected t-statistics are printed in parentheses and ***, **, * corresponds to the 1%, 5%, and 10% significance level. Due to data availability, the sample period is limited to 1990 - 2021 for the CBOE volatility index and to 1972 - 2021 for all other state variables.

Node	VIX	NBER	Liquidity	Sentiment
BP: Quantiles (main)	0.47***	0.90***	−0.33***	0.05
	(6.92)	(4.61)	(−5.90)	(0.42)
Weighting scheme	0.40***	0.54***	−0.24***	0.06
	(5.60)	(2.70)	(−4.63)	(0.60)
Positive earnings	0.15	−0.43**	−0.07	0.25**
	(1.45)	(−2.44)	(−0.99)	(2.40)
Size restriction	0.34***	0.84***	−0.32***	−0.01
	(5.62)	(3.25)	(−5.59)	(−0.13)
Sorting variable lag	0.37***	0.89***	−0.26***	0.02
	(4.91)	(4.64)	(−5.31)	(0.17)
BP: Exchanges	0.33***	0.31	−0.22***	0.32***
	(3.14)	(1.56)	(−3.39)	(2.90)
Financials	0.45***	0.67	−0.16**	0.14*
	(3.60)	(1.32)	(−2.54)	(1.85)
Double sort	0.46***	1.06***	−0.30***	−0.14
	(9.83)	(3.85)	(−4.37)	(−1.13)
BP: Quantiles (second)	0.43***	0.78***	−0.33***	−0.10
	(8.41)	(2.86)	(−4.37)	(−0.75)
Rebalancing	0.15**	0.51***	−0.17***	0.01
	(2.24)	(2.71)	(−3.19)	(0.07)
Utilities	0.29***	0.79***	−0.29***	−0.06
	(6.85)	(3.68)	(−3.73)	(−0.69)
Stock-age restriction	0.15	0.32	−0.21**	0.27
	(1.47)	(0.98)	(−2.37)	(1.46)
Price restriction	0.08	0.24	−0.12	−0.35***
	(0.71)	(0.69)	(−1.61)	(−2.72)
Positive book equity	0.24***	−0.13	−0.01	0.10
	(3.37)	(−0.43)	(−0.20)	(0.74)

In order to investigate for which group of sorting variables the relations with our economic indicators are most pronounced, we implement the following panel regression:

$$\text{Residual MAD}^n_{v,t} = \alpha^n_g + \beta^n_1 \cdot \text{Economic indicator}_t + \beta^n_2 \cdot \text{Economic indicator}_t \cdot \mathbb{1}_{v \in g} + \epsilon^n_t.$$

For each node n, we regress the residual mean absolute difference of each sorting variable v on the economic indicator and the economic indicator interacted with a dummy variable that takes a value of one if the sorting variable belongs to the group g under investigation. To facilitate comparisons, we run this panel regressions for each node n repeatedly, including only one interaction term for each of the eight groups of sorting variables. The interaction coefficients for these regressions in Table VII.1,

Table VII.2, and Table VII.3 in the Internet Appendix VII, show that the relations in Table 6 for market volatility, recessions, and liquidity are more pronounced for sorting variables belonging to the groups size, trading frictions, and profitability. Contrary, these relations tend to be notably weaker for sorting variables belonging to the groups financing and investments. The positive coefficients in Table 6 for sentiment and the decision nodes "positive earnings" and "breakpoint exchanges" tend to be also more pronounced for sorting variables belonging to the groups "size" and "trading frictions" (Table VII.4 in the Internet Appendix VII).

In conclusion, non-standard errors in portfolio sorts are countercyclical, i.e., they are higher when stock markets are volatile, in a recession, or illiquid. Our results imply that one needs to be careful when relating premia from portfolio sorts to either other premia or any explanatory variables over time because the non-standard errors constitute a measurement error that is not just pure noise but correlated to the right-hand side variable, leading to biased estimates. For instance, our results show a strong relation between mean absolute differences and NBER recessions. If a researcher wants to understand if a certain premium is higher in recessions, it is not clear if an estimated positive coefficient is driven by the underlying economic relation or rather by the correlation of the recession indicator and the measurement error induced by non-standard errors. To address this concern, we recommend analyzing if the explanatory variable is systematically related to non-standard errors in portfolio sorts. Conditionally on answering the previous question with yes, investigating whether the proposed relation to the explanatory variable is robust to possible specifications of the premium can alleviate concerns about a potentially spurious relation.

6 Conclusion

We analyze the impact of seemingly innocuous methodological choices (non-standard errors) in a somewhat standardized procedure in asset pricing, namely portfolio sorts.

First, we find that methodological choices in portfolio sorts have a significant impact on the size, significance, and monotonicity of estimated premia. This methodology-induced uncertainty about the premium is larger than the well-understood uncertainty induced by repeatedly sampling from the same population (i.e., standard errors). Although methodological choices induce considerable variation in estimated premia, we find that – irrespective of the specification – at least 90% of all estimated premia are positive, and 58% to 74% are statistically significant at the 5% level. This finding alleviates concerns in the recent literature about a "replication crisis" suggesting that a large amount of published premia does not exist and is limited to specific methodological choices ("p-hacking"). In fact, our evidence suggests that the data-generating process does feature most premia in our sample, which are remarkably robust to a large set of methodological choices. This holds in particular for those premia that have strong theoretical underpinnings.

Second, while size-related decision nodes such as a “size restriction”, “NYSE breakpoints”, or the “weighting scheme” as suggested by Hou et al. (2020) do induce large variation in estimated premia, other decision nodes are even more important. In particular, the time lag between information arrival and portfolio formation and the exclusion of stocks with negative earnings introduce at least as much variation as size-related decision nodes. Even when we take out the size-related decision nodes, the average non-standard error is still significant and only reduced by around 30%. Moreover, this reduction is not homogeneous across groups of sorting variables and introduces large differences in the monotonicity of premia. This has two important implications: Seemingly innocuous decisions beyond size-related decisions have large impacts on the size, significance, and monotonicity of premia. Moreover, fixing size-related decision nodes is hardly enough to control for non-standard errors in portfolio sorts. Therefore, we encourage investigating the distribution of premia conditional on all methodological choices, which can be done conveniently using the code we provide online.

Third, non-standard errors in portfolio sorts show strong time-variation driven by market volatility, the business cycle, and episodes of illiquidity. Thereby, non-standard errors introduce correlated measurement error. This can bias coefficient estimates when relating premia to (predictor) variables such as business cycle indicators, which are correlated with the measurement error in premia. Therefore, non-standard errors in portfolio sorts might have an impact beyond the question of whether there is a premium but also on research investigating the follow-up question about the economic sources of the premium.

References

Abarbanell, J. S. and B. J. Bushee (1998): "Abnormal returns to a fundamental analysis strategy," *The Accounting Review*, 19–45.

Allaire, J., Y. Xie, J. McPherson, J. Luraschi, K. Ushey, A. Atkins, H. Wickham, J. Cheng, W. Chang, and R. Iannone (2023): *rmarkdown: Dynamic Documents for R*, r package version 2.20.

Altman, E. I. (1968): "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy," *The Journal of Finance*, 23, 589–609.

Amihud, Y. (2002): "Illiquidity and stock returns: cross-section and time-series effects," *Journal of Financial Markets*, 5, 31–56.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2006): "The cross-section of volatility and expected returns," *The Journal of Finance*, 61, 259–299.

Areal, N. (2021): *frenchdata: Download Data Sets from Kenneth's French Finance Data Library Site*, r package version 0.2.0.

Asness, C. and A. Frazzini (2013): "The devil in HML's details," *The Journal of Portfolio Management*, 39, 49–68.

Auguie, B. (2017): *gridExtra: Miscellaneous Functions for "Grid" Graphics*, r package version 2.3.

Baker, M. and J. Wurgler (2006): "Investor sentiment and the cross-section of stock returns," *The Journal of Finance*, 61, 1645–1680.

Balakrishnan, K., E. Bartov, and L. Faurel (2010): "Post loss/profit announcement drift," *Journal of Accounting and Economics*, 50, 20–41.

Bali, T. G., N. Cakici, and R. F. Whitelaw (2011): "Maxing out: Stocks as lotteries and the cross-section of expected returns," *Journal of Financial Economics*, 99, 427–446.

Bali, T. G., R. F. Engle, and S. Murray (2016): *Empirical asset pricing: The cross section of stock returns*, John Wiley & Sons.

Ball, R., J. Gerakos, J. T. Linnainmaa, and V. Nikolaev (2016): "Accruals, cash flows, and operating profitability in the cross section of stock returns," *Journal of Financial Economics*, 121, 28–45.

Banz, R. W. (1981): "The relationship between return and market value of common stocks," *Journal of Financial Economics*, 9, 3–18.

Banz, R. W. and W. J. Breen (1986): "Sample-dependent results using accounting and market data: some evidence," *The Journal of Finance*, 41, 779–793.

Barbee Jr, W. C., S. Mukherji, and G. A. Raines (1996): "Do sales–price and debt–equity explain stock returns better than book–market and firm size?" *Financial Analysts Journal*, 52, 56–60.

Basu, S. (1983): "The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence," *Journal of Financial Economics*, 12, 129–156.

Belo, F., X. Lin, and S. Bazdresch (2014): "Labor hiring, investment, and stock return predictability in the cross section," *Journal of Political Economy*, 122, 129–177.

Bhandari, L. C. (1988): "Debt/equity ratio and expected common stock returns: Empirical evidence," *The Fournal of Finance*, 43, 507–528.

Blitz, D., J. Huij, and M. Martens (2011): "Residual momentum," *Journal of Empirical Finance*, 18, 506–521.

Boudoukh, J., R. Michaely, M. Richardson, and M. R. Roberts (2007): "On the importance of measuring payout yield: Implications for empirical asset pricing," *The Journal of Finance*, 62, 877–915.

Bradshaw, M. T., S. A. Richardson, and R. G. Sloan (2006): "The relation between corporate financing activities, analysts' forecasts and stock returns," *Journal of Accounting and Economics*, 42, 53–85.

Brennan, M. J., T. Chordia, and A. Subrahmanyam (1998): "Alternative factor specifications, security characteristics, and the cross-section of expected stock returns," *Journal of Financial Economics*, 49, 345–373.

Chan, L. K., N. Jegadeesh, and J. Lakonishok (1996): "Momentum strategies," *The Journal of Finance*, 51, 1681–1713.

Chan, L. K., J. Lakonishok, and T. Sougiannis (2001): "The stock market valuation of research and development expenditures," *The Journal of Finance*, 56, 2431–2456.

Chen, A. Y. (2022): "Most claimed statistical findings in cross-sectional return predictability are likely true," *arXiv preprint arXiv:2206.15365*.

Chen, A. Y. and T. Zimmermann (2022): "Open Source Cross-Sectional Asset Pricing," *Critical Finance Review*, 27, 207–264.

Chordia, T., A. Goyal, and A. Saretto (2020): "Anomalies and false rejections," *The Review of Financial Studies*, 33, 2134–2179.

Cochrane, J. H. (2011): "Presidential address: Discount rates," *The Journal of Finance*, 66, 1047–1108.

Cooper, M. J., H. Gulen, and M. J. Schill (2008): "Asset growth and the cross-section of stock returns," *The Journal of Finance*, 63, 1609–1651.

Coqueret, G. (2022): "Forking paths in empirical studies," *Available at SSRN 3999379*.

Dahl, D. B., D. Scott, C. Roosen, A. Magnusson, and J. Swinton (2019): *xtable: Export Tables to LaTeX or HTML*, r package version 1.8-4.

Daniel, K. and S. Titman (2006): "Market reactions to tangible and intangible information," *The Journal of Finance*, 61, 1605–1643.

Datar, V. T., N. Y. Naik, and R. Radcliffe (1998): "Liquidity and stock returns: An alternative test," *Journal of Financial Markets*, 1, 203–219.

Davis, J. L., E. F. Fama, and K. R. French (2000): "Characteristics, covariances, and average returns: 1929 to 1997," *The Journal of Finance*, 55, 389–406.

De Bondt, W. F. and R. Thaler (1985): "Does the stock market overreact?" *The Journal of Finance*, 40, 793–805.

Desai, H., S. Rajgopal, and M. Venkatachalam (2004): "Value-glamour and accruals mispricing: One anomaly or two?" *The Accounting Review*, 79, 355–385.

Dichev, I. D. (1998): "Is the risk of bankruptcy a systematic risk?" *The Journal of Finance*, 53, 1131–1147.

Dickey, D. A. and W. A. Fuller (1979): "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American statistical association*, 74, 427–431.

Fama, E. F. and K. R. French (1992): "The cross-section of expected stock returns," *The Journal of Finance*, 47, 427–465.

——— (1993): "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 33, 3–56.

——— (1996): "The CAPM is wanted, dead or alive," *The Journal of Finance*, 51, 1947–1958.

——— (2015): "A five-factor asset pricing model," *Journal of Financial Economics*, 116, 1–22.

Fama, E. F. and J. D. MacBeth (1973): "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy*, 81, 607–636.

Feng, G., S. Giglio, and D. Xiu (2020): "Taming the factor zoo: A test of new factors," *The Journal of Finance*, 75, 1327–1370.

Firke, S. (2023): *janitor: Simple Tools for Examining and Cleaning Dirty Data*, r package version 2.2.0.

Foster, G., C. Olsen, and T. Shevlin (1984): "Earnings releases, anomalies, and the behavior of security returns," *The Accounting Review*, 574–603.

Francis, J., R. LaFond, P. M. Olsson, and K. Schipper (2004): "Costs of equity and earnings attributes," *The Accounting Review*, 79, 967–1010.

Frazzini, A. and L. H. Pedersen (2014): "Betting against beta," *Journal of Financial Economics*, 111, 1–25.

Gao, X. and J. R. Ritter (2010): "The marketing of seasoned equity offerings," *Journal of Financial Economics*, 97, 33–52.

George, T. J. and C.-Y. Hwang (2004): "The 52-week high and momentum investing," *The Journal of Finance*, 59, 2145–2176.

Green, J., J. R. Hand, and X. F. Zhang (2013): "The supraview of return predictive signals," *Review of Accounting Studies*, 18, 692–730.

Grolemund, G. and H. Wickham (2011): "Dates and Times Made Easy with lubridate," *Journal of Statistical Software*, 40, 1–25.

Guenther, D. A. and A. J. Rosman (1994): "Differences between COMPUSTAT and CRSP SIC codes and related effects on research," *Journal of Accounting and Economics*, 18, 115–128.

Hafzalla, N., R. Lundholm, and E. Matthew Van Winkle (2011): "Percent accruals," *The Accounting Review*, 86, 209–236.

Hahn, J. and H. Lee (2009): "Financial constraints, debt capacity, and the cross-section of stock returns," *The Journal of Finance*, 64, 891–921.

HARVEY, C. R. (2017): "Presidential address: The scientific outlook in financial economics," *The Journal of Finance*, 72, 1399–1440.

——— (2019): "Replication in financial economics," *Available at SSRN 3409466.*

HARVEY, C. R., Y. LIU, AND H. ZHU (2016): ". . . and the cross-section of expected returns," *The Review of Financial Studies*, 29, 5–68.

HASLER, M. (2021): "Is the Value Premium Smaller Than We Thought?" *Available at SSRN 3886984.*

——— (2022): "Looking Under the Hood of Data-Mining," *Available at SSRN 4279944.*

HAUGEN, R. A. AND N. L. BAKER (1996): "Commonality in the determinants of expected stock returns," *Journal of Financial Economics*, 41, 401–439.

HIRSHLEIFER, D., K. HOU, S. H. TEOH, AND Y. ZHANG (2004): "Do investors overvalue firms with bloated balance sheets?" *Journal of Accounting and Economics*, 38, 297–331.

HOU, K., H. MO, C. XUE, AND L. ZHANG (2021): "An augmented q-factor model with expected growth," *Review of Finance*, 25, 1–41.

HOU, K., C. XUE, AND L. ZHANG (2014): "Digesting Anomalies: An Investment Approach," *The Review of Financial Studies*, 28, 650–705.

——— (2020): "Replicating anomalies," *The Review of Financial Studies*, 33, 2019–2133.

HRIBAR, P. AND D. W. COLLINS (2002): "Errors in estimating accruals: Implications for empirical research," *Journal of Accounting Research*, 40, 105–134.

HUANG, A. G. (2009): "The cross section of cashflow volatility and expected stock returns," *Journal of Empirical Finance*, 16, 409–429.

JEGADEESH, N. (1990): "Evidence of predictable behavior of security returns," *The Journal of Finance*, 45, 881–898.

JEGADEESH, N. AND J. LIVNAT (2006): "Revenue surprises and stock returns," *Journal of Accounting and Economics*, 41, 147–171.

JENSEN, T. I., B. T. KELLY, AND L. H. PEDERSEN (2021): "Is there a replication crisis in finance?" Tech. rep., National Bureau of Economic Research.

KAHLE, K. M. AND R. A. WALKLING (1996): "The impact of industry classifications on financial research," *Journal of financial and quantitative analysis*, 31, 309–335.

KAMIŃSKI, B., M. JAKUBCZYK, AND P. SZUFEL (2018): "A framework for sensitivity analysis of decision trees," *Central European Journal of Operations Research*, 26, 135–159.

KIM, J. H. AND P. I. JI (2015): "Significance testing in empirical finance: A critical review and assessment," *Journal of Empirical Finance*, 34, 1–14.

KOMSTA, L. AND F. NOVOMESTKY (2022): *moments: Moments, Cumulants, Skewness, Kurtosis and Related Tests*, r package version 0.14.1.

KUHN, M., S. JACKSON, AND J. CIMENTADA (2022): *corrr: Correlations in R*, r package version 0.4.4.

KÖSTLMEIER, S. (2019): *monotonicity: Test for Monotonicity in Expected Asset Returns, Sorted by Portfolios*, r package version 1.3.1.

Lakonishok, J., A. Shleifer, and R. W. Vishny (1994): "Contrarian investment, extrapolation, and risk," *The Journal of Finance*, 49, 1541–1578.

Lamont, O., C. Polk, and J. Saá-Requejo (2001): "New evidence on measuring financial constraints: moving beyond the KZ index," *The RAND Journal of Economics*, 32, 101–128.

Lev, B. and D. Nissim (2004): "Taxable income, future earnings, and equity values," *The Accounting Review*, 79, 1039–1074.

Linnainmaa, J. T. and M. R. Roberts (2018): "The history of the cross-section of stock returns," *The Review of Financial Studies*, 31, 2606–2649.

Luo, H., I. Liu, and N. Tripathy (2021): "A study on firms with negative book value of equity," *International Review of Finance*, 21, 145–182.

Lyandres, E., L. Sun, and L. Zhang (2008): "The new issues puzzle: Testing the investment-based explanation," *The Review of Financial Studies*, 21, 2825–2855.

McLean, R. D. and J. Pontiff (2016): "Does academic research destroy stock return predictability?" *The Journal of Finance*, 71, 5–32.

Menkveld, A. J., A. Dreber, F. Holzmeister, J. Huber, M. Johannesson, M. Kirchler, S. Neusüss, M. Razen, U. Weitzel, and Others (2022): "Non-standard errors," *The Journal of Finance (forthcoming).*

Mitton, T. (2022): "Methodological variation in empirical corporate finance," *The Review of Financial Studies*, 35, 527–575.

Morey, M. R. and S. Yadav (2018): "Documentation of the file drawer problem in academic finance journals," *The Journal of Investing*, 27, 143–147.

Müller, K. and H. Wickham (2022): *tibble: Simple Data Frames*, r package version 3.1.8.

Müller, K., H. Wickham, D. A. James, and S. Falcon (2023): *RSQLite: SQLite Interface for R*, r package version 2.3.0.

Newey, W. K. and K. D. West (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

——— (1994): "Automatic lag selection in covariance matrix estimation," *The Review of Economic Studies*, 61, 631–653.

Novy-Marx, R. (2011): "Operating leverage," *Review of Finance*, 15, 103–134.

——— (2013): "The other side of value: The gross profitability premium," *Journal of Financial Economics*, 108, 1–28.

Ohlson, J. A. (1980): "Financial ratios and the probabilistic prediction of bankruptcy," *Journal of Accounting Research*, 109–131.

Ooms, J. (2014): "The jsonlite Package: A Practical and Consistent Mapping Between JSON Data and R Objects," *arXiv:1403.2805 [stat.CO].*

Pástor, L. and R. F. Stambaugh (2003): "Liquidity risk and expected stock returns," *Journal of Political Economy*, 111, 642–685.

Patton, A. J. and A. Timmermann (2010): "Monotonicity in asset returns: New tests with applications to the term structure, the CAPM, and portfolio sorts," *Journal of Financial Economics*, 98, 605–625.

Penman, S. H., S. A. Richardson, and I. Tuna (2007): "The book-to-price effect in stock returns: accounting for leverage," *Journal of Accounting Research*, 45, 427–467.

R Core Team (2022): *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

R Special Interest Group on Databases (R-SIG-DB), H. Wickham, and K. Müller (2022): *DBI: R Database Interface*, r package version 1.1.3.

Richardson, S. A., R. G. Sloan, M. T. Soliman, and I. Tuna (2005): "Accrual reliability, earnings persistence and stock prices," *Journal of Accounting and Economics*, 39, 437–485.

Robinson, D., A. Hayes, and S. Couch (2023): *broom: Convert Statistical Objects into Tidy Tibbles*, r package version 1.0.3.

Rosenthal, R. (1979): "The file drawer problem and tolerance for null results." *Psychological Bulletin*, 86.

Scheuch, C., S. Voigt, and P. Weiss (2023): *Tidy Finance with R*, Chapman and Hall/CRC, 1st ed.

Scholz, F. and A. Zhu (2019): *kSamples: K-Sample Rank Tests and their Combinations*, r package version 1.2-9.

Scholz, F. W. and M. A. Stephens (1987): "K-sample Anderson–Darling tests," *Journal of the American Statistical Association*, 82, 918–924.

Sharpsteen, C. and C. Bracken (2023): *tikzDevice: R Graphics Output in LaTeX Format*, r package version 0.12.4.

Shumway, T. (1997): "The delisting bias in CRSP data," *The Journal of Finance*, 52, 327–340.

Sloan, R. G. (1996): "Do stock prices fully reflect information in accruals and cash flows about future earnings?" *The Accounting Review*, 289–315.

Soebhag, A., B. van Vliet, and P. Verwijmeren (2023): "Non-Standard Errors in Asset Pricing: Mind your sorts," *Available at SSRN.*

Soliman, M. T. (2008): "The use of DuPont analysis by market participants," *The Accounting Review*, 83, 823–853.

Thomas, J. and F. X. Zhang (2011): "Tax expense momentum," *Journal of Accounting Research*, 49, 791–821.

Thomas, J. K. and H. Zhang (2002): "Inventory changes and future returns," *Review of Accounting Studies*, 7, 163–187.

Titman, S., K. J. Wei, and F. Xie (2004): "Capital investments and stock returns," *Journal of Financial and Quantitative Analysis*, 39, 677–700.

Tuzel, S. (2010): "Corporate real estate holdings and the cross-section of stock returns," *The Review of Financial Studies*, 23, 2268–2302.

Ushey, K. (2022): *renv: Project Environments*, r package version 0.16.0.

Vaughan, D. (2022): *slider: Sliding Window Functions*, r package version 0.3.0.

Vaughan, D. and M. Dancho (2022): *furrr: Apply Mapping Functions in Parallel using Futures*, r package version 0.3.1.

Whited, T. M. and G. Wu (2006): "Financial constraints risk," *The Review of Financial Studies*, 19, 531–559.

Wickham, H. (2016): *ggplot2: Elegant Graphics for Data Analysis*, Springer-Verlag New York.

——— (2022): *stringr: Simple, Consistent Wrappers for Common String Operations*, r package version 1.5.0.

——— (2023): *forcats: Tools for Working with Categorical Variables (Factors)*, r package version 1.0.0.

Wickham, H., M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani (2019): "Welcome to the tidyverse," *Journal of Open Source Software*, 4, 1686.

Wickham, H. and J. Bryan (2023): *readxl: Read Excel Files*, r package version 1.4.2.

Wickham, H., R. François, L. Henry, K. Müller, and D. Vaughan (2023a): *dplyr: A Grammar of Data Manipulation*, r package version 1.1.0.

Wickham, H., M. Girlich, and E. Ruiz (2023b): *dbplyr: A 'dplyr' Back End for Databases*, r package version 2.3.1.

Wickham, H. and L. Henry (2023): *purrr: Functional Programming Tools*, r package version 1.0.1.

Wickham, H., J. Hester, and J. Bryan (2023c): *readr: Read Rectangular Text Data*, r package version 2.1.4.

Wickham, H., J. Ooms, and K. Müller (2023d): *RPostgres: Rcpp Interface to PostgreSQL*, r package version 1.4.5.

Wickham, H. and D. Seidel (2022): *scales: Scale Functions for Visualization*, r package version 1.2.1.

Wickham, H., D. Vaughan, and M. Girlich (2023e): *tidyr: Tidy Messy Data*, r package version 1.3.0.

Xie, Y. (2023): *knitr: A General-Purpose Package for Dynamic Report Generation in R*, r package version 1.42.

Xie, Y., J. Cheng, and X. Tan (2023): *DT: A Wrapper of the JavaScript Library 'DataTables'*, r package version 0.27.

Xing, Y. (2008): "Interpreting the value effect through the Q-theory: An empirical investigation," *The Review of Financial Studies*, 21, 1767–1795.

Zeileis, A. and G. Grothendieck (2005): "zoo: S3 Infrastructure for Regular and Irregular Time Series," *Journal of Statistical Software*, 14, 1–27.

Zeileis, A. and T. Hothorn (2002): "Diagnostic Checking in Regression Relationships," *R News*, 2, 7–10.

Zeileis, A., S. Köll, and N. Graham (2020): "Various Versatile Variances: An Object-Oriented Implementation of Clustered Covariances in R," *Journal of Statistical Software*, 95, 1–36.

A Construction of sorting variables

In this section we describe all details to construct the 68 sorting variables, which we analyze in this paper. Note that we replace any negative values of total assets (At), sales (SALE), capital expenditures (CAPX), and inventories (INVT) as missing.

A.1 Financing

Composite debt issuance. We follow Lyandres et al. (2008) and measure composite debt issuance (CDI) for each firm in each fiscal year t from Compustat annual data as the logarithmic growth rate in the book value of debt from fiscal year $t-5$ to fiscal year t. The book value of debt is measured by the sum of current debt (DLC) and long-term debt (DLTT).

Composite share issuance. Daniel and Titman (2006) propose to measure composite share issuance (CSI) from CRSP data as the difference between the change in market equity and the cumulative log return of a stock. Both, the change in market equity and cumulative log returns are measured in each month from year t to year $t-5$.

Change in common equity. To capture the change in common equity (DBE) according to Richardson et al. (2005), we calculate the following ratio from Compustat annual data:

$$DBE_t = \frac{CEQ_t - CEQ_{t-1}}{AT_{t-1}},$$

where CEQ represents common equity and AT total assets.

Change in current operating liabilities. Richardson et al. (2005) measure the change in current operating liabilities (DCOL) for each firm in each fiscal year t from annual Compustat data:

$$DCOL_t = \frac{(LCT_t - DLC_t) - (LCT_{t-1} - DLC_{t-1})}{AT_{t-1}},$$

where LCT are current liabilities, DLC short-term debt, and AT total assets. We replace missing values of DLC with zero.

Change in financial liabilities. We define the change in financial liabilities (DFNL) similar to Richardson et al. (2005) for each stock in each fiscal year t from annual Compustat data in the following way:

$$DFNL_t = \frac{(DLTT_t + DLC_t + PSTK_t) - (DLTT_{t-1} + DLC_{t-1} + PSTK_{t-1})}{AT_{t-1}},$$

where DLTT is long-term debt, DLC short-term debt, PSTK the value of preferred stocks, and AT total assets. Missing values of DLTT, DLC, and PSTK are set to zero if at least one of the three variables is available.

Net debt financing. We follow Bradshaw et al. (2006) and compute net debt financing (NDF) for each stock in each fiscal year t from annual Compustat data as:

$$NDF_t = \frac{DLTIS_t - DLTR_t + DLCCH_t}{\frac{1}{2}\big(AT_t + AT_{t-1}\big)},$$

where DLTIS are cash proceeds from the issuance of long-term debt, DLTR are cash payments for long-term debt reductions, DLCCH are the net changes in current debt, and AT total assets. We replace missing values of DLCCH with zero. Data starts in January 1972 due to data availability of financing variables.

Net equity financing. We measure net equity financing (NEF) similar to Bradshaw et al. (2006) for each stock in each fiscal year t from annual Compustat data:

$$NEF_t = \frac{SSTK_t - PRSTKC_t - DV_t}{\frac{1}{2}\big(AT_t + AT_{t-1}\big)},$$

where SSTK are proceeds from the sale of common and preferred stocks, PRSTKC are payments for the repurchase of common and preferred stocks, DV are cash payments for dividends, and AT total assets. Data starts in January 1972 due to data availability of financing variables.

Net external financing. We capture net external financing (NXF) for each stock in each fiscal year t similar to Bradshaw et al. (2006) by the sum of net debt financing and net equity financing. Both variables are described above. Data starts in January 1972 due to data availability of financing variables.

A.2 Intangibles

Advertisement expenses to market equity. Chan et al. (2001) suggest measuring the advertising expense to market ratio (ADM) as advertising expenses (Compustat item XAD) divided by market equity, which is obtained from CRSP at the end of each fiscal year. We exclude observations with negative advertising expenses. We start our measure in January 1973 to ensure sufficient data coverage.

Cash-flow volatility. We follow Huang (2009) and compute operating cash flows to sales for each stock in each fiscal quarter q from quarterly Compustat data:

$$Operating\ cash\ flows_q = \frac{IBQ_q + DPQ_q + (WCAP_q - WCAP_{q-1})}{SALEQ_q},$$

where IBQ are quarterly income before extraordinary items, DPQ are quarterly depreciation and amortizations, WCAPQ are quarterly working capital, and SALEQ are quarterly sales. Cash-flow volatility (CFV) for each stock in each fiscal quarter corresponds to the standard deviation of operating cash flows during the past 16 quarters. We require a minimum of eight observations. We start our measure in January 1978 to ensure sufficient data coverage.

Earnings' predictability. Francis et al. (2004) define split-adjusted earnings per share (EPSA) from Compustat data as earnings per share (EPSPX) divided by the adjustment factor (AJEX). We follow Francis et al. (2004) and measure earnings predictability (EPRD) for each stock as the residual volatility (u_t) from the following auto-regressive process:

$$EPSA_t = \alpha + \beta \cdot EPSA_{t-1} + u_t.$$

Moreover, we measure this auto-regressive process over the last ten years and always require ten years of non-missing observations.

Hiring rate. We follow Belo et al. (2014) and obtain the hiring rate (HR) for each stock in each fiscal year t from annual Compustat data:

$$HR_t = \frac{EMP_t - EMP_{t-1}}{\frac{1}{2}\big(EMP_t + EMP_{t-1}\big)},$$

where EMP represents the number of employees. Moreover, we exclude firms with a hiring rate of zero.

Kaplan and Zingales index for financing constraints. We obtain the Kaplan and Zingales index (KZI) for each firm in each fiscal year from annual Compustat data by following Lamont et al. (2001):

$$\begin{aligned} KZI_t = &- 1.002 \cdot \frac{IB_t + DP_t}{PPENT_{t-1}} + 0.283 \cdot \frac{AT_t + ME_t - CEQ_t - TXDB_t}{AT_t} + 3.139 \cdot \frac{DLC_t + DLTT_t}{DLC_t + DLTT_t + SEQ_t} \\ &- 39.368 \cdot \frac{DVC_t + DVP_t}{PPENT_{t-1}} - 1.315 \cdot \frac{CHE_t}{PPENT_{t-1}}, \end{aligned}$$

where IB corresponds to income before extraordinary items, DP to depreciation and amortization, PPENT to property, plant, and equipment, AT to total assets, ME to market equity from CRSP at the end of each fiscal year, CEQ to common equity, TXDB to deferred taxes, DLC to current debt, DLTT to long-term debt, SEQ to shareholder equity, DVC to dividends of common stock, DVP to dividends of preferred stocks, and CHE to cash holdings.

Labor force efficiency. We define the labor force efficiency (LFE) as in Abarbanell and Bushee (1998) for each firm in each fiscal year t from annual Compustat data:

$$LFE_t = \left(\frac{SALE_t}{EMP_t} - \frac{SALE_{t-1}}{EMP_{t-1}}\right) \Big/ \frac{SALE_{t-1}}{EMP_{t-1}},$$

where SALE corresponds to sales and EMP to employees.

Operating leverage. We follow Novy-Marx (2013) and compute operating leverage (OL) from Compustat data as cost of goods sold (COGS) plus selling, general and administrative expenses (XSGA), both scaled by current total assets (AT).

R&D expenses to market equity. Chan et al. (2001) propose to compute the R&D expense to market ratio (RDM) as R & D expenses (Compustat item XRD) divided by market equity from the end of each fiscal year. We obtain market equity data from CRSP and include only observations with positive R & D expenses. We start our measure in January 1976 because R &D expenses were standardized in 1975.

Real-estate ratio. We define the real-estate ratio (RER) similar to Tuzel (2010) with Compustat data. Prior to 1983, it corresponds to the sum of buildings (PPENB) and capital leases (PPENLS) scaled by net property, plant and equipment (PPENT). After the end of 1983, it is measured as the sum of buildings at cost (FATB) and leases at cost (FATL), both divided by gross property, plant and equipment (PPEGT). Subsequently, we winsorize the real estate ratios in each fiscal year at the 1 % and 99 % percentile. The industry-adjusted real-estate ratio is obtained by subtracting the industry average real-estate ratio from each stock-specific real-estate ratio. We use 2-digit SIC codes to assign stocks to industries. We always require at least five observations to calculate the industry average each year. Note that real estate data starts in 1969, limiting the observation period for this specific sorting variable. Data for the real estate ratio starts in January 1970 due to data availability.

Tangibility. We capture the tangibility (TAN) of each firm in each fiscal year according to Hahn and Lee (2009) from annual Compustat data:

$$TAN_t = \frac{CHE_t}{AT_t} + 0.715 \cdot \frac{RECT_t}{AT_t} + 0.547 \cdot \frac{INVT_t}{AT_t} + 0.535 \cdot \frac{PPEGT_t}{AT_t},$$

where CHE corresponds to cash holdings, RECT to accounts receivable, INVT to inventory, PPEGT to property, plant, and equipment, and AT to total assets.

Whited and Wu index for financing constraints. We closely follow Whited and Wu (2006) and measure financing constraints for each firm in each fiscal year t from annual Compustat data:

$$\begin{aligned} WW_t = & -0.091 \cdot \frac{IB_t + DP_t}{AT_t} - 0.062 \cdot \mathbb{1}_{DVPSX\_F>0} + 0.021 \cdot \frac{DLTT_t}{AT_t} - 0.044 \cdot \ln\left(AT_t\right) + 0.102 \cdot ISG_t \\ & -0.035 \cdot \frac{SALE_t - SALE_{t-1}}{SALE_{t-1}}, \end{aligned}$$

where IB is income before extraordinary items, DP depreciation and amortization, AT total assets, $\mathbb{1}_{DVPSX>0}$ a dummy variable equal to one if the firm pays out cash dividends (DVPSX_F), DLTT long-term debt and SALE sales. Moreover, ISG is the industry growth rate of sales, while industries are defined by 3-digit SIC codes. Industries with less than two firms are excluded. Since Whited and Wu (2006) estimate this index with quarterly data, we replace the annual growth rates in industry sales growth and stock-specific sales growth with their implied quarterly compounded growth rates. Lastly, we winsorize the distribution of each sub-variable of the Whited and Wu index at the 1% and 99% quantile.

A.3 Investments

Abnormal corporate investment. We measure abnormal corporate investments (ACI) from Compustat annual data for each firm in each fiscal year t as in Titman et al. (2004):

$$ACI_t = \frac{CE_t}{\frac{1}{3}\left(CE_{t-1} + CE_{t-2} + CE_{t-3}\right)} - 1,$$

where CE corresponds to capital expenditures (Compustat item CAPX) divided by sales (SALE). We follow Hou et al. (2020) and exclude stocks with sales below 10 million dollars.

Asset growth. We follow Cooper et al. (2008) and measure asset growth (AG) for each stock in each fiscal year t from Compustat data as the change in total assets (AT) from year t to year $t-1$, divided by total assets from year $t-1$.

Change in net operating assets We measure net operating assets for each stock in each fiscal year t from Compustat annual data:

$$Net\ operating\ assets_t = (AT_t - CHE_t - IVAO_t) - (AT_t - DLC_t - DLTT_t - MIB_t - PSTK_t - CEQ_t),$$

where AT corresponds to total assets, CHE to cash and short-term investments, IVAO to other investments and advances, DLC to current liabilities, DLTT to long-term debt, MIB to minority interests, PSTK to the value of preferred stocks, and CEQ to common equity. Missing values in DLC, DLTT, MIB, and PSTK are set to zero. The change in net operating assets (DNOA) is then the difference between net operating assets of fiscal year t and fiscal year $t-1$ scaled by total assets of year $t-1$.

Change in property, plant, equipment and inventory to assets. We add the annual change in gross property, plant and equipment (PPEGT) to the annual change in inventory (INVT) and scale this sum by one-year-lagged total assets (AT). Thus, we obtain the change in property, plant, equipment and inventories (DPIA) as in Lyandres et al. (2008) from Compustat data.

Change in net non-cash working capital. Following Richardson et al. (2005) we define non-cash working capital from Compustat data as:

$$WC_t = (ACT_t - CHE_t) - (LCT_t - DLC_t),$$

where ACT corresponds to current assets, CHE to cash, LCT to current liabilities, and DLC to short-term debt. We set missing values of DLC to zero. The change in net non-cash working capital (DWC) corresponds to the change of WC from fiscal year t to fiscal year $t-1$ scaled by total assets from fiscal year $t-1$.

Investment growth. We compute investment growth (IG) from annual Compustat data as the annual change in capital expenditures (CAPX) from fiscal year t to year $t-1$, scaled by capital expenditures from year $t-1$.

Inventory changes. Thomas and Zhang (2002) suggest measuring the change in inventory (DINV) from Compustat data as the annual change in inventories (INVT) from fiscal year t to fiscal year $t-1$, divided by average total assets (AT) over the fiscal year t and $t-1$.

Net operating assets. We compute net operating assets (NOA) from Compustat data:

$$NOA_t = \frac{(AT_t - CHE_t - IVAO_t) - (AT_t - DLC_t - DLTT_t - MIB_t - PSTK_t - CEQ_t)}{AT_{t-1}},$$

where AT is total assets, CHE cash and short-term investments, IVAO other investments and advances, DLC short-term debt, DLTT long-term debt, MIB minority interest, PSTK preferred stock, and CEQ common equity. We replace missing values of DLC, DLTT, MIB, and PSTK as zero.

Operating accruals. The definition of operating accruals (OA) before 1988 closely follows Sloan (1996):

$$OA_t = \frac{(\Delta ACT_t - \Delta CHE_t) - (\Delta LCT_t - \Delta DLC_t - \Delta TXP_t) - DP_t}{AT_{t-1}},$$

where ACT is current assets, CHE cash, LCT current liabilities, DLC short-term debt, TXP taxes payable, and DP depreciation and amortization. Moreover, we replace missing values of DLC and TXP with zero. Due to data availability, we follow Hribar and Collins (2002) and compute operating accruals from 1988 and onward as:

$$OA_t = \frac{NI_t - OANCF_t}{AT_{t-1}},$$

where NI is net income and OANCF corresponds to net cash flow from operations. All items are from Compustat data.

Percent total accruals. Hafzalla et al. (2011) suggest measuring percent total accruals (PTA) from Compustat data as total accruals (TA) divided by the absolute value of net income (NI). Before 1988 we follow Hou et al. (2020) and define PTA as:

$$\begin{aligned} PTA_t =& \big(\Delta(ACT_t - CHE_t - LCT_t + DLC_t) + \Delta(AT_t - ACT_t - IVAO_t - LT_t + LCT_t + DLTT_t) \\ &+ \Delta(IVST_t + IVAO_t - DLTT_t - DLC_t - PSTK_t)\big)/|NI_t|, \end{aligned}$$

where ACT is current assets, LCT current liabilities, DLC short-term debt, AT total assets, IVAO investments and advances, LT total liabilities, DLTT long-term debt, IVST short-term investments, PSTK preferred stock, and NI net income. Δ refers to the change from fiscal year t to fiscal year $t-1$. Moreover, missing values of IVAO, DLTT, DLC, IVST, and PSTK are set to zero. From 1988 and, thereafter, we follow Hribar and Collins (2002) and measure PTA from Compustat data as

$$PTA_t = \frac{NI_t - OANCF_t - IVNCF_t - FINCF_t + SSTK_t - PRSTKC_t - DV_t}{|NI_t|},$$

where NI corresponds to net income, OANCF to total operating cash flows, IVNCF to total investing cash flows, FINCF to total financing cash flows, SSTK to the sale of stocks, PRSTKC to stock repurchases, and DV to dividends. Moreover, we set missing valus of SSTK and DV to zero.

A.4 Momentum

Cumulative abnormal returns around earnings' announcements. We follow Chan et al. (1996) and estimate abnormal returns around quarterly earnings' announcements in month t as the difference between the individual stock return $r_{i,d}$ and the market index $r_{m,d}$ on day d. We cumulate these abnormal returns around the following 4-day event window including two days before the quarterly earnings announcement, the day of the announcement, and one day after:

$$ABR_{i,t} = \sum_{d=-2}^{d=1} (r_{i,d} - r_{m,d}).$$

The quarterly earnings announcement date corresponds to Compustat item RDQ and has to be after the fiscal quarter end date to exclude potential recording errors. Data starts in January 1972 because the earnings announcement date RDQ is only available from 1972 onwards.

Return momentum. We compute return momentum (MOM) for each stock in each month as the cumulative return from month $t-12$ to month $t-2$ skipping the most recent month as in Fama and French (1996).

Residual momentum. As in Blitz et al. (2011), we define the 11-month residual momentum (RMOM) in each month and for each stock as cumulative residual returns from month $t-2$ to month $t-12$, scaled by the standard deviation of residual returns over the same time horizon. Residual returns are obtained in each month from regressing monthly excess stock returns from month $t-1$ to month $t-36$ on the Fama and French (1993) three-factor model. Throughout these rolling regressions, we always require 36 monthly returns.

Revenue surprise. Similar to Jegadeesh and Livnat (2006), we construct revenues per share from Compustat quarterly data for each stock i in each quarter q as:

$$Revenues\ per\ share_q = \frac{SALEQ_q}{CSHPRQ_q \cdot AJEXQ_q},$$

where SALEQ corresponds to quarterly revenues, CHSPRQ to the correction factor for quarterly shares outstanding, and AJEXQ to quarterly shares outstanding. Revenue surprises (RS) correspond then to the change in revenues per share over the last four quarters scaled by the standard deviation of the change in revenues per share over the last eight quarters. We require at least six quarterly observations for this rolling standard deviation. Lastly, the earnings announcement date has to be after the fiscal quarter end date. Data starts in January 1972 because the earnings announcement date RDQ is only available from 1972 onwards.

Standardized unexpected earnings. As in Foster et al. (1984), we calculate unexpected earnings from quarterly Compustat data for each stock in each quarter q as the change in split-adjusted earnings per share from its value four quarters ago:

$$Unexpected\ earnings\ per\ share_q = \frac{EPSPXQ_q}{AJEXQ_q} - \frac{EPSPXQ_{q-4}}{AJEXQ_{q-4}},$$

where ESPSPXQ are quarterly earnings per share and AJEXQ denotes the number of shares outstanding in each quarter. Then, standardized unexpected earnings (SUE) are defined as unexpected earnings per share divided by the standard deviation of unexpected earnings per share over the previous eight quarters. We require at least six quarterly observations for this rolling standard deviation. Moreover, the earnings announcement date has to be before the fiscal quarter end date. Data starts in January 1972 because the earnings announcement date RDQ is only available from 1972 onwards.

Tax expense surprise. We follow Thomas and Zhang (2011), and calculate tax expense surprises (TES) for each stock in each quarter q as the change in tax expenses per share over the last four quarters scaled by assets per share from four quarters ago ($q-4$):

$$TES_q = \frac{\frac{TXTQ_q}{CSHPRQ_q \cdot AJEXQ_q} - \frac{TXTQ_{q-4}}{CSHPRQ_{q-4} \cdot AJEXQ_{q-4}}}{\frac{ATQ_q}{CSHPRQ_q \cdot AJEXQ_q}},$$

where TXTQ represents quarterly tax expenses, ATQ quarterly total assets, AJEXQ quarterly shares outstanding, and CSHPRQ the factor to adjust quarterly shares outstanding. We exclude firms that do not pay taxes from our sample and require the earnings announcement date to be after the fiscal quarter end date. We follow Hou et al. (2020) and start our calculation in January 1976 to ensure data availability of this measure.

52-week high. We define the 52-week high (52W), similar to George and Hwang (2004), for each stock in each month t as the daily split-adjusted stock price at the end of each month scaled by the highest daily split-adjusted stock price over the previous 12 months.

A.5 Profitability

Asset turnover. We follow Soliman (2008) and compute asset turnover (ATO) from Compustat data as sales (SALE) divided by net operating assets from the previous fiscal year:

$$ATO_t = \frac{SALE_t}{(AT_{t-1} - CHE_{t-1} - IVAO_{t-1}) - (AT_{t-1} - DLC_{t-1} - DLTT_{t-1} - MIB_{t-1} - PSTK_{t-1} - CEQ_{t-1})},$$

where Compustat item AT are total assets, item CHE are cash and short-term investments, and IVAO are other investments and advances. Moreover, item DLC represents debt in current liabilities, DLTT long-term debt, MIB minority interests, PSTK preferred stocks, and CEQ common equity. We follow Hou et al. (2020) and replace missing values of IVAO, DLC, DLTT, MIB, and PSTK with zero. Similar to Hou et al. (2020), we exclude observations with negative net operating assets.

Book leverage. Similar to Fama and French (1992), we compute the book leverage (BL) of each firm in each fiscal year by the ratio of total assets (Compustat item AT) and book equity. The definition of book equity follows from Davis et al. (2000) and is disclosed below when describing the book-to-market ratio.

Cash-based operating profitability. The definition of cash-based operating profitability (CBOP) closely follows Ball et al. (2016) and is based on Compustat data:

$$CBOP_t = \frac{REVT_t - COGS_t - XSGA_t + XRD_t - \Delta RECT_t - \Delta INVT_t - \Delta XPP_t + \Delta DRC_t + \Delta DRLT_t + \Delta AP_t + \Delta XACC_t}{AT_t},$$

where REVT is total revenue, COGS are cost of goods sold, XSGA are selling, general and administrative expenses, and XRD are R&D expenses. Moreover, $\Delta RECT_t$ is the change in accounts receivable, $\Delta INVT$ the change in inventory, ΔXPP is the change in prepaid expenses, $\Delta DRC_t + \Delta DRLT$ the change in deferred revenues, ΔAP the change in trade accounts payable and $\Delta XACC$ is the change in accrued expenses. We follow Hou et al. (2020) and set missing values of XRD and all missing changes to zero.

Capital turnover. We measure capital turnover (CTO) from Compustat data as sales (SALE) divided by total assets (AT) from the previous fiscal year (Haugen and Baker (1996)).

Gross profits to assets. We follow Novy-Marx (2013) and obtain gross profits to assets (GPA) from Compustat data as total revenues (REVT) minus cost of goods sold (COGS), scaled by current total assets (AT).

Ohlson's O-score. Ohlson (1980) suggests evaluating the financial stability of a firm with the following linear relation:

$$\begin{aligned} O_t = &-1.32 - 0.407 \cdot \log(AT_t) + 6.03 \cdot \frac{DLC_t + DLTT_t}{AT_t} - 1.43 \cdot \frac{ACT_t - LCT_t}{AT_t} + 0.076 \cdot \frac{LCT_t}{AT_t} - 1.72 \cdot \mathbb{1}_{LT_t > AT_t} - 2.37 \cdot \frac{NI_t}{AT_t} \\ &- 1.83 \cdot \frac{PI_t + DP_t}{LT_t} + 0.285 \cdot \mathbb{1}_{NI_t<0 \ \& \ NI_{t-1}<0} - 0.521 \cdot \frac{NI_t - NI_{t-1}}{|NI_t| + |NI_{t-1}|}. \end{aligned}$$

All data items are obtained from Compustat: AT corresponds to total assets, DLC to short-term debt, DLTT to long-term debt, ACT to current assets, LCT to current liabilities, LT to total liabilities, PI to pretax income, DP to depreciation and amortization, and NI to net income. We follow Hou et al. (2020) and winsorize all variables except for dummy variables at the 1% and 99 % quantile of their respective distribution.

Operating profits to book equity. We closely follow Fama and French (2015) and compute operating profits to book equity for each firm in each fiscal year t from Compustat annual data:

$$OPE_t = \frac{REVT_t - COGS_t - XSGA_t - XINT_t}{BE_t},$$

where REVT corresponds to total revenues, COGS to cost of goods sold, XSGA to selling, general and administrative expenses XINT to interest expenses, and BE to book equity. The definition of book equity follows the disclosed definition for the variable book-to-market (below). Moreover, missing values in COGS, XSGA, and XINT are set to zero.

Return on assets. We obtain data on return on assets (ROA) for each stock in each fiscal quarter q from Compustat quarterly data and closely follow Balakrishnan et al. (2010):

$$ROA_q = \frac{IBQ_q}{ATQ_{q-1}},$$

where IBQ corresponds to quarterly income before extraordinary items, and ATQ represents quarterly total assets. Moreover, the earnings announcement date of each record has to be after the fiscal quarter end date to ensure consistent recording. Data starts in January 1972 because the earnings announcement date RDQ is only available from 1972 onwards.

Return on equity. Hou et al. (2014) define return on equity (ROE) for each firm in each fiscal quarter as:

$$ROE_q = \frac{IBQ_q}{BEQ_{q-1}},$$

where IBQ corresponds to quarterly income before extraordinary items and BEQ to quarterly book equity. Quarterly book equity (BEQ) is computed as the book equity of shareholders plus balance sheet deferred taxes and investment tax credit minus the book value of preferred stock. Depending on data availability, we measure shareholders' equity by SEQQ, or the sum of common equity (CEQQ) and the par value of preferred stock (PSTKQ), or if all previous items are unavailable by total assets (ATQ) minus total liabilities (LTQ). The book value of preferred stocks corresponds to PSTKRQ, to PSTKQ if PSTKRQ is unavailable or to zero if both are unavailable. Balance sheet deferred taxes and investment tax credit is TXDITCQ, TXDBQ if TXDITCQ is missing or zero if both are missing. Data starts in January 1972 because the earnings announcement date RDQ is only available from 1972 onwards.

Taxable income to book income. We closely follow Green et al. (2013) and compute taxable income to book income (TBI) for each firm in each fiscal year t from Compustat annual data as:

$$TBI_t = \frac{PI_t}{NI_t},$$

where PI is pretax income and NI is net income. Moreover, we require positive pretax and net income.

Altman's Z-score. We measure the Altman (1968) Z-score for each firm in each fiscal year from Compustat annual data by the following definition:

$$Z_t = 1.2 \cdot \frac{ACT_t - LCT_t}{AT_t} + 1.4 \cdot \frac{RE_t}{AT_t} + 3.3 \cdot \frac{OIADP_t}{AT_t} + 0.6 \cdot \frac{ME_t}{LT_t} + \frac{SALE_t}{AT_t},$$

where ACT is current assets, LCT current liabilities, AT total assets, RE retained earnings, OIADP earnings before interest and taxes, ME market equity from the end of the fiscal year (from CRSP), LT total liabilities, and SALE corresponds to sales. Lastly, we winsorize the distributions of all five sub-variables of the Z-score at the 1% and 99% quantile in each fiscal year.

A.6 Size

Size. We follow Fama and French (1992) and compute the size of each stock (ME) in each month as the natural logarithm of the market equity. We obtain market equity data from CRSP by multiplying the shares outstanding (SHROUT) with the corresponding share price (PRC).

A.7 Trading frictions

Amihud illiquidity measure. Amihud (2002) proposes to measure the illiquidity of each firm on each day d from daily CRSP data as the absolute daily return scaled by the daily dollar trading volume:

$$return\, to\, volume_d = \frac{|RET_d|}{PRC_d * VOL_d},$$

where RET is the daily return, PRC is the daily price, and VOL is the daily volume of stocks traded. The Amihud illiquidity measure (AMI) for each firm in each month t corresponds to the average return to volume estimate over the last six months. We require at least 50 observations for this average and adjust the trading volume of NASDAQ stocks according to Gao and Ritter (2010).

Beta relative to the market. We compute the market beta (BETA) for each stock in each month t from monthly CRSP data and similar as in Fama and MacBeth (1973). Specifically, we run the following time series regression over the previous five years:

$$r_t^e = \alpha + \beta_1 \cdot (MKT_t - R_t^f) + u_t.$$

Moreover, we require at least 24 monthly observations for the regression above. The market beta for each firm in each month t corresponds to the regression coefficient β_1. Data on the market factor MKT_t is obtained from Kenneth French's website.

Frazzini and Pedersen beta. Frazzini and Pedersen (2014) suggest measuring the beta (BFP) for each stock i and each month t from daily CRSP data as:

$$BFP_{i,t} = \frac{\hat{\rho} \cdot \hat{\sigma_i}}{\hat{\sigma_m}},$$

where $\hat{\sigma_i}$ corresponds to the standard deviation of each stock i measured as the standard deviation of daily logarithmic returns over the previous 750 days. Moreover, $\hat{\sigma_m}$ is the standard deviation of the market and is obtained as the standard deviation of daily logarithmic returns over the previous 750 days. Throughout the calculations of these standard deviations for each month t, we require at least 120 daily observations. Lastly, $\hat{\rho}$ is the return correlation between the market m and stock i. We estimate this return correlation for each month t over the last 750 daily returns. When estimating the return correlation, we use overlapping 3-day logarithmic returns for each stock i on each day d: $r_{i,d} = \sum_{k=-2}^{0} \ln(1 + r_{i,d+k})$ instead of one-day raw returns.

Dollar trading volume. We follow Brennan et al. (1998) and compute dollar trading volume (DTV) from daily CRSP data as the average dollar trading volume from month $t-1$ to month $t-6$. We require at least 50 days of observations when computing this average. Dollar trading volume is defined as share price (PRC) multiplied by the number of shares traded on each day (VOL). Moreover, we adjust dollar trading volume from NASDAQ according to Gao and Ritter (2010).

Idiosyncratic skewness relative to the Fama and French (1993) model. We regress the daily excess returns of each stock on the Fama and French (1993) factor model:

$$r_t^e = \alpha + \beta_1 \cdot (MKT_t - R_t^f) + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t + u_t.$$

Throughout these regressions, we require at least 15 daily observations for each month. Idiosyncratic skewness (ISKEW) relative to the Fama and French (1993) model is then measured in each month as the skewness of residuals u_t (Bali et al. (2016)).

Idiosyncratic volatility relative to the Fama and French (1993) model. We follow Ang et al. (2006) and compute idiosyncratic volatility relative to the Fama and French (1993) factor model (IVOL) as the volatility of residuals from the following regression:

$$r_t^e = \alpha + \beta_1 \cdot (MKT_t - R_t^f) + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t + u_t.$$

In detail, we regress in each month the excess return of each stock on the Fama and French (1993) factor model using daily returns from CRSP and Kenneth French. Moreover, we require at least 15 daily observations for each month.

Maximum daily return. We compute the maximum daily return (MDR) for each stock in each month t similar to Bali et al. (2011) from daily CRSP data as the maximal daily return in each month t. Moreover, we require at least 15 return observations for each month t.

Short-term reversal. We follow Jegadeesh (1990) and measure the short-term reversal (SREV) for each firm in each month t from monthly CRSP data as the stock return during month t. We require a valid return on month t for all stocks. All sorting variables are subsequently lagged according to decision node "sorting variable lag", i.e., this definition does not produce a look-ahead bias.

Share turnover. Datar et al. (1998) propose to measure the daily share turnover (TUR) of each stock on each day as the number of shares traded (VOL) scaled by the number of shares outstanding (SHROUT) on the same day. The variable share turnover for each firm in each month t is then the average daily share turnover over the previous six months. Throughout this calculation, we require at least 50 daily observations. Lastly, we adjust the trading volume of NASDAQ stocks according to Gao and Ritter (2010).

A.8 Valuation

Assets to market equity. Similar to Fama and French (1992) we measure assets to market equity (AM) for each stock in each fiscal year by total assets (AT) divided by market equity (CRSP) from the end of the fiscal year t. We exclude observations with negative total assets.

Book equity to market equity. This paper follows Davis et al. (2000) and computes the book-to-market ratio (BM) as book equity from Compustat divided by market equity from CRSP. Market equity is measured at the end of each fiscal year. Book equity corresponds to the book equity of shareholders plus balance sheet deferred taxes and investment tax credit (Compustat item TXDITC or TXDB + ITCB if TXDITC is unavailable) minus the book value of preferred stock. Depending on data availability, we measure shareholders' equity by SEQ, or the sum of common equity (CEQ) and the par value of preferred stock (PSTK), or if all previous items are unavailable by total assets (AT) minus total liabilities (LT). The book value of preferred stock corresponds in the following order either to the redemption value (PSTKRV), or the liquidation value (PSTKL), or if all previous items are unavailable to the par value (PSTK). We replace missing values of TXDITC or the book value of preferred stock with zero.

Cash-flow to market equity. Lakonishok et al. (1994) suggest measuring the cash-flow-to-price ratio (CFM) from Compustat data as income before extraordinary items (IB) plus depreciation (DP), both divided by market equity (CRSP) from the end of the fiscal year. We exclude all stocks with negative cash flows.

Debt to market equity. Following Bhandari (1988), the debt to market ratio (DM) is defined as short-term debt (Compustat item DLC) plus long-term debt (Compustat item DLTT) divided by market equity obtained from CRSP at the end of each fiscal year. We exclude stocks with missing DLC and DLTT observations.

Enterprise book equity to market equity. We follow Penman et al. (2007) and obtain enterprise book equity scaled by market equity (EBM) for each firm in each fiscal year t as net debt plus book equity scaled by net debt plus market equity:

$$EBM_t = \frac{(DLTT_t + DLC_t + PSTK_t + DVPA_t - TSTKP_t) - CHE_t + BE_t}{(DLTT_t + DLC_t + PSTK_t + DVPA_t - TSTKP_t) - CHE_t + ME_t},$$

where DLTT corresponds to long-term debt, DLC to current liabilities, PSTK to the value of preferred stock, DVPA to preferred dividends in arrears, TSTKP to preferred treasury stock, CHE to cash and short-term investments, and ME to the market equity from CRSP measured at the end of each fiscal year t. Lastly, book equity BE is computed as common equity (CEQ) plus TSTKP minus DVPA. Lastly, missing observations in DVPA and TSTKP are set to zero. We require that the sum of net debt and book equity as well as the sum of net debt plus market equity are positive.

Earnings to market equity. We follow Basu (1983) and compute the earnings-to-price ratio (EM) as income before extraordinary items (Compustat item IB) divided by market equity from CRSP. Market equity corresponds to the end of each fiscal year. We exclude firms with negative earnings.

Net debt to market equity. Similar to Penman et al. (2007), net debt to price (NDM) is measured from Compustat annual data for each stock in each fiscal year t in the following way:

$$NDM_t = \frac{(DLTT_t + DLC_t + PSTK_t + DVPA_t - TSTKP_t) - CHE_t}{ME_t},$$

where DLTT corresponds to long term-debt, DLC to current liabilities, PSTK to the value of preferred stock, DVPA to preferred dividends in arrears, TSTKP to preferred treasury stock, CHE to cash and short-term investments, and ME to the market equity from CRSP measured at the end of each fiscal year t. Lastly, missing observations in DVPA and TSTKP are set to zero.

Net payout yield. Boudoukh et al. (2007) suggest measuring the net payout yield (NPY) of each stock in the following way:

$$NPY_t = \frac{(DVC_t + PRSTKC_t + \Delta PSTKRV_t \cdot \mathbb{1}_{\Delta PSTKRV<0}) - (SSTK_t - \Delta PSTKRV_t \cdot \mathbb{1}_{\Delta PSTKRV>0})}{ME_t},$$

where DVC are dividends from common stock, PRSTKC is the purchase of common and preferred stock, PSTKRV is the value of the net number of preferred stocks outstanding, and SSTK reflects the sale of common and preferred stocks. $\mathbb{1}_{\Delta PSTKRV<0}$ is a dummy variable that has value one if the annual change in PSTKRV is negative and zero otherwise. Market equity (ME) is from CRSP and corresponds to the end of each fiscal year. Moreover, we exclude stocks with negative net payouts. Data starts in January 1972 because of sufficient data coverage for the sale of common and preferred stocks.

Operating cash-flow to market equity. We follow Desai et al. (2004) and compute the ratio of operating cash-flows to price (OCM) as operating cash flows from Compustat divided by the market equity at the end of each fiscal year from CRSP. Before 1988, we measure operating cash flows as funds from operations (FOPT) minus the change in working capital (item WCAP) and as net cash flows from operating activities (OANCF) thereafter. Moreover, we exclude firms with negative operating cash flows. Data starts in January 1972 because of sufficient data coverage for funds from operations.

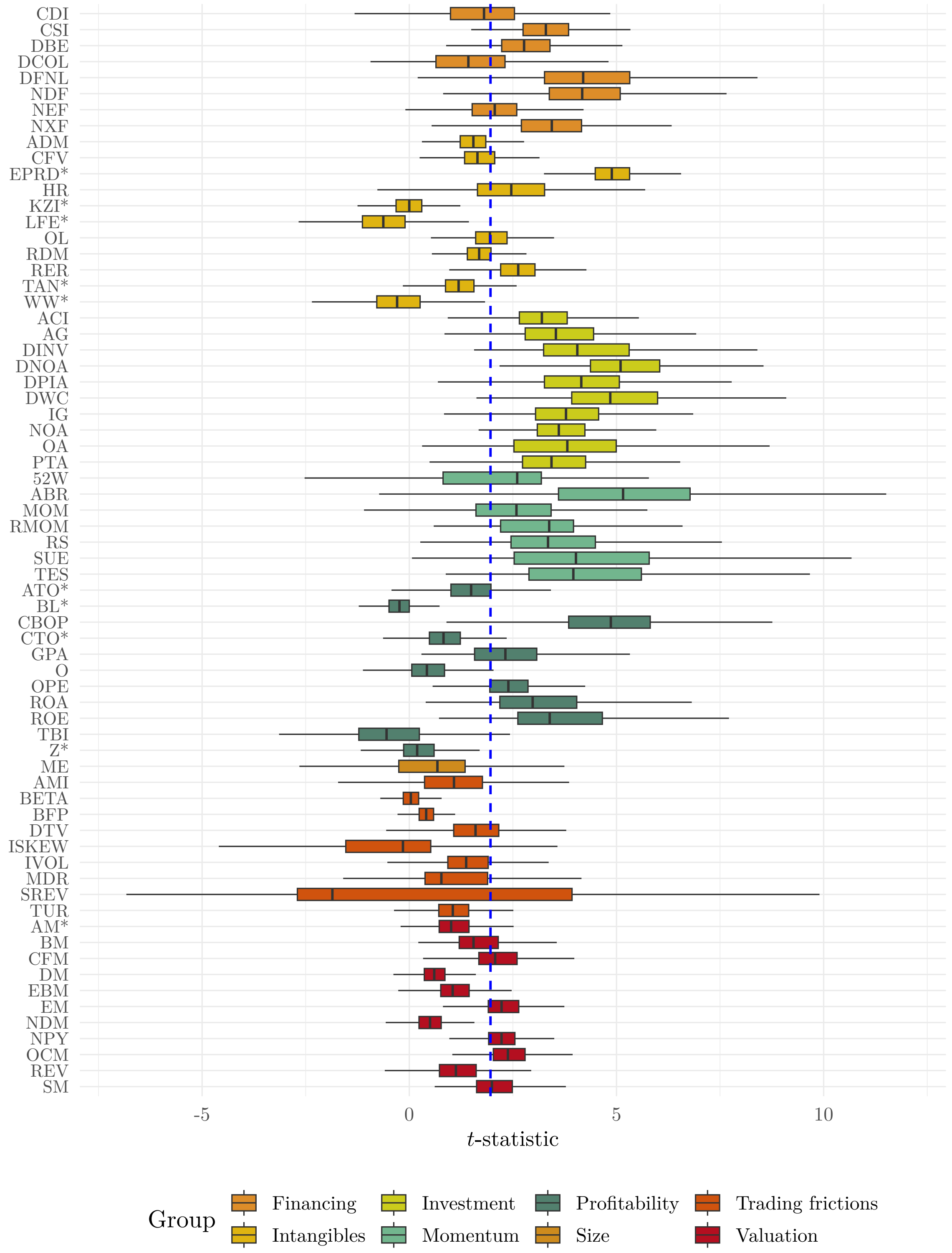
Long-term reversal. We measure the long-term reversal effect (REV) suggested by De Bondt and Thaler (1985) for each stock in each month t by the cumulative returns from month $t-60$ to month $t-13$.

Sales to market equity We compute the sales to price ratio (SM) as sales (Compustat item SALE) divided by the market equity at the end of each fiscal year (Barbee Jr et al., 1996). Stocks with negative sales are excluded.

B Distribution of t-statistics

In Figure B.1, we show the distribution of t-statistics for all sorting variables. We show additional t-statistics for intercepts of three asset pricing models (i.e., the CAPM, FF5, and Q5) in the Internet Appendix III.

Figure B.1: Variation in t-statistics across sorting variables.
This figure shows the estimated t-statistics in box plots for all sorting variables across all decision nodes. The vertical axis shows the associated sorting variable, while the color scheme connects each sorting variable to the respective category. A t-value of 1.96 is indicated by the vertical dashed line.



C Impact of decision nodes

Table C.1: Impact of decision node: Breakpoint quantiles (main)

For each branch of node "breakpoint quantiles (main)", we show the mean statistics across sorting variables within each group in separate panels. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: Financing

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
5	0.28	0.13	(0.01, 0.06)	1.06	0.82	3.84	0.99	0.74	0.49
10	0.36	0.18	(0.02, 0.06)	1.16	0.63	3.44	0.97	0.77	0.34

Panel B: Intangibles

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
5	0.20	0.10	(0.00, 0.01)	0.63	0.53	3.81	0.84	0.36	0.35
10	0.25	0.14	(0.00, 0.01)	0.74	0.55	4.09	0.80	0.35	0.25

Panel C: Investment

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
5	0.35	0.15	(0.02, 0.09)	1.24	0.78	3.46	1.00	0.95	0.67
10	0.47	0.19	(0.02, 0.08)	1.22	0.71	3.46	1.00	0.97	0.43

Panel D: Momentum

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
5	0.42	0.34	(0.09, 0.16)	1.99	0.58	2.90	0.98	0.81	0.70
10	0.54	0.47	(0.10, 0.15)	2.04	0.45	2.83	0.98	0.82	0.59

Panel E: Profitability

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
5	0.20	0.16	(0.01, 0.04)	0.88	0.61	3.59	0.81	0.39	0.40
10	0.26	0.19	(0.01, 0.04)	0.91	0.57	3.45	0.82	0.41	0.30

Panel F: Size

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
5	0.06	0.23	(0.10, 0.04)	1.39	-0.24	5.55	0.68	0.07	0.17
10	0.12	0.28	(0.11, 0.06)	1.55	1.41	9.50	0.71	0.10	0.12

Panel G: Trading frictions

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
5	0.14	0.23	(0.02, 0.07)	1.17	0.56	3.95	0.79	0.16	0.13
10	0.16	0.28	(0.02, 0.07)	1.17	0.62	4.69	0.76	0.16	0.10

Panel H: Valuation

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
5	0.25	0.12	(0.00, 0.01)	0.57	0.85	4.25	0.98	0.37	0.52
10	0.30	0.15	(0.00, 0.01)	0.61	0.77	3.93	0.99	0.35	0.37

Table C.2: Impact of decision node: Weighting scheme

For each branch of node "weighting scheme", we show the mean statistics across sorting variables within each group in separate panels. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: Financing

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
EW	0.36	0.17	(0.03, 0.07)	1.21	0.90	3.87	1.00	0.87	0.53
VW	0.28	0.15	(0.01, 0.05)	1.04	0.74	3.86	0.96	0.63	0.44

Panel B: Intangibles

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
EW	0.24	0.13	(0.00, 0.01)	0.72	0.62	4.03	0.82	0.40	0.35
VW	0.21	0.14	(0.00, 0.01)	0.75	0.55	4.37	0.82	0.30	0.34

Panel C: Investment

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
EW	0.44	0.19	(0.08, 0.10)	1.47	0.72	3.49	1.00	0.99	0.67
VW	0.38	0.18	(0.03, 0.07)	1.24	0.72	3.35	1.00	0.93	0.68

Panel D: Momentum

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
EW	0.50	0.43	(0.15, 0.17)	2.22	0.47	2.75	0.98	0.84	0.73
VW	0.46	0.36	(0.07, 0.16)	1.92	0.77	3.50	0.99	0.79	0.68

Panel E: Profitability

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
EW	0.24	0.18	(0.03, 0.05)	0.98	0.64	3.57	0.82	0.42	0.38
VW	0.22	0.17	(0.01, 0.04)	0.87	0.75	3.98	0.81	0.38	0.41

Panel F: Size

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
EW	0.11	0.26	(0.11, 0.07)	1.71	1.43	9.96	0.70	0.12	0.16
VW	0.08	0.25	(0.10, 0.02)	1.28	0.02	6.08	0.70	0.05	0.17

Panel G: Trading frictions

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
EW	0.17	0.28	(0.03, 0.08)	1.30	0.73	4.45	0.77	0.20	0.15
VW	0.13	0.23	(0.02, 0.06)	1.02	0.33	4.09	0.77	0.12	0.11

Panel H: Valuation

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
EW	0.31	0.14	(0.00, 0.01)	0.64	0.89	4.22	0.99	0.46	0.59
VW	0.24	0.11	(0.00, 0.00)	0.53	0.88	4.03	0.98	0.26	0.46

Table C.3: Impact of decision node: Positive earnings filter

For each branch of node "positive earnings filter", we show the mean statistics across sorting variables within each group in separate panels. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: Financing

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.36	0.20	(0.02, 0.07)	1.23	0.57	3.31	0.99	0.78	0.48
Yes	0.28	0.15	(0.01, 0.05)	1.06	0.42	3.26	0.98	0.72	0.49

Panel B: Intangibles

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.25	0.15	(0.01, 0.01)	0.78	0.55	3.96	0.83	0.37	0.33
Yes	0.20	0.12	(0.00, 0.01)	0.66	0.18	3.37	0.80	0.33	0.36

Panel C: Investment

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.42	0.20	(0.07, 0.09)	1.35	0.68	3.24	1.00	0.96	0.57
Yes	0.41	0.18	(0.05, 0.09)	1.32	0.64	3.19	1.00	0.96	0.78

Panel D: Momentum

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.51	0.39	(0.10, 0.15)	2.02	0.57	3.14	0.99	0.83	0.71
Yes	0.45	0.39	(0.11, 0.17)	2.13	0.60	2.90	0.97	0.80	0.69

Panel E: Profitability

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.28	0.18	(0.01, 0.03)	0.87	0.62	3.62	0.83	0.44	0.42
Yes	0.19	0.15	(0.00, 0.05)	0.84	0.56	3.59	0.80	0.36	0.38

Panel F: Size

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.09	0.22	(0.09, 0.05)	1.52	1.45	11.08	0.70	0.07	0.11
Yes	0.10	0.29	(0.12, 0.04)	1.45	0.14	4.24	0.69	0.10	0.22

Panel G: Trading frictions

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.16	0.26	(0.02, 0.07)	1.15	0.62	4.96	0.79	0.15	0.12
Yes	0.13	0.26	(0.03, 0.07)	1.19	0.52	3.70	0.76	0.17	0.14

Panel H: Valuation

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.31	0.15	(0.00, 0.01)	0.60	0.86	4.12	1.00	0.38	0.54
Yes	0.24	0.12	(0.00, 0.01)	0.55	0.82	4.17	0.97	0.34	0.51

Table C.4: Impact of decision node: Size restriction

For each branch of node "size restriction", we show the mean statistics across sorting variables within each group in separate panels. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: Financing

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.37	0.21	(0.05, 0.09)	1.38	0.56	3.18	0.99	0.81	0.48
0.2	0.27	0.13	(0.00, 0.02)	0.92	0.40	3.26	0.97	0.67	0.51

Panel B: Intangibles

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.25	0.15	(0.01, 0.02)	0.87	0.55	3.96	0.82	0.43	0.38
0.2	0.20	0.11	(0.00, 0.00)	0.61	0.22	3.73	0.81	0.26	0.32

Panel C: Investment

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.46	0.22	(0.09, 0.12)	1.54	0.61	3.17	1.00	0.97	0.70
0.2	0.37	0.16	(0.03, 0.06)	1.17	0.61	3.04	1.00	0.95	0.71

Panel D: Momentum

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.54	0.48	(0.15, 0.21)	2.39	0.40	2.55	0.98	0.83	0.73
0.2	0.41	0.34	(0.05, 0.09)	1.56	0.25	2.14	0.98	0.78	0.69

Panel E: Profitability

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.26	0.21	(0.03, 0.07)	1.06	0.56	3.35	0.84	0.41	0.40
0.2	0.19	0.14	(0.00, 0.01)	0.69	0.42	3.05	0.79	0.36	0.38

Panel F: Size

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.19	0.32	(0.12, 0.12)	1.76	1.25	7.31	0.75	0.23	0.21
0.2	0.07	0.24	(0.13, 0.00)	1.21	-0.74	3.23	0.71	0.02	0.18

Panel G: Trading frictions

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.17	0.31	(0.03, 0.10)	1.43	0.57	3.98	0.74	0.21	0.16
0.2	0.13	0.22	(0.01, 0.04)	0.90	0.12	3.30	0.80	0.12	0.11

Panel H: Valuation

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.32	0.17	(0.00, 0.01)	0.72	0.79	3.83	0.99	0.48	0.60
0.2	0.22	0.09	(0.00, 0.00)	0.41	0.49	3.33	0.97	0.21	0.43

Table C.5: Impact of decision node: Sorting variable lag

For each branch of node "sorting variable lag", we show the mean statistics across sorting variables within each group in separate panels. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: Financing

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
1m	0.47	0.20	(0.00, 0.01)	0.95	0.57	2.94	1.00	0.99	0.43
3m	0.34	0.18	(0.04, 0.06)	1.20	0.65	3.75	0.98	0.78	0.54
6m	0.33	0.16	(0.02, 0.05)	1.15	0.61	3.77	0.99	0.78	0.50
FF	0.26	0.14	(0.00, 0.07)	1.13	0.73	4.04	0.97	0.67	0.42

Panel B: Intangibles

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
3m	0.23	0.14	(0.01, 0.01)	0.80	0.64	4.65	0.83	0.36	0.34
6m	0.22	0.13	(0.01, 0.01)	0.76	0.58	4.37	0.81	0.34	0.34
FF	0.22	0.13	(0.00, 0.01)	0.77	0.59	4.26	0.80	0.36	0.35

Panel C: Investment

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
3m	0.47	0.20	(0.08, 0.09)	1.43	0.51	3.17	1.00	0.98	0.71
6m	0.42	0.18	(0.05, 0.07)	1.29	0.63	3.44	1.00	0.97	0.68
FF	0.35	0.15	(0.02, 0.06)	1.14	0.74	3.54	1.00	0.94	0.62

Panel D: Momentum

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
1m	0.66	0.29	(0.06, 0.03)	1.07	0.21	3.44	0.92	0.75	0.61
3m	0.55	0.24	(0.06, 0.11)	1.44	0.42	2.97	1.00	0.94	0.84
6m	0.32	0.13	(0.01, 0.03)	0.82	0.14	4.06	1.00	0.69	0.59

Panel E: Profitability

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
3m	0.27	0.18	(0.03, 0.04)	0.96	0.63	3.68	0.81	0.43	0.44
6m	0.20	0.15	(0.01, 0.02)	0.79	0.57	3.63	0.81	0.38	0.37
FF	0.16	0.15	(0.00, 0.02)	0.76	0.59	3.58	0.78	0.28	0.28

Panel F: Size

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
1m	0.22	0.20	(0.13, 0.07)	1.64	3.15	17.75	0.96	0.17	0.30
3m	-0.04	0.33	(0.11, 0.05)	1.46	0.61	6.48	0.41	0.04	0.05
6m	0.10	0.24	(0.02, 0.02)	1.04	0.47	6.77	0.71	0.05	0.16

Panel G: Trading frictions

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
1m	0.27	0.17	(0.03, 0.04)	0.97	0.56	5.14	0.86	0.29	0.20
3m	0.10	0.14	(0.01, 0.01)	0.72	0.32	4.87	0.78	0.11	0.10
6m	0.08	0.13	(0.00, 0.01)	0.66	0.19	4.73	0.68	0.08	0.10

Panel H: Valuation

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
1m	0.24	0.15	(0.00, 0.02)	0.79	1.39	6.82	0.99	0.23	0.57
3m	0.28	0.14	(0.00, 0.01)	0.65	0.99	4.73	0.98	0.36	0.53
6m	0.27	0.13	(0.00, 0.01)	0.59	0.92	4.43	0.98	0.36	0.51
FF	0.28	0.13	(0.00, 0.00)	0.58	0.77	3.62	1.00	0.37	0.53

Table C.6: Impact of decision node: Breakpoint exchanges
For each branch of node "breakpoint exchanges", we show the mean statistics across sorting variables within each group in separate panels. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: Financing

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
All	0.36	0.19	(0.03, 0.06)	1.22	0.46	3.25	0.98	0.79	0.48
NYSE	0.28	0.14	(0.01, 0.05)	1.01	0.62	3.90	0.98	0.71	0.49

Panel B: Intangibles

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
All	0.24	0.15	(0.01, 0.01)	0.82	0.51	4.04	0.81	0.38	0.35
NYSE	0.21	0.12	(0.00, 0.01)	0.67	0.52	4.02	0.83	0.32	0.34

Panel C: Investment

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
All	0.45	0.21	(0.07, 0.09)	1.40	0.56	3.13	1.00	0.96	0.66
NYSE	0.37	0.16	(0.04, 0.07)	1.24	0.68	3.47	1.00	0.96	0.69

Panel D: Momentum

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
All	0.53	0.45	(0.13, 0.18)	2.24	0.50	2.75	0.99	0.83	0.72
NYSE	0.44	0.36	(0.07, 0.14)	1.84	0.46	2.65	0.98	0.80	0.68

Panel E: Profitability

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
All	0.25	0.20	(0.03, 0.05)	0.98	0.60	3.40	0.82	0.41	0.39
NYSE	0.21	0.15	(0.01, 0.03)	0.81	0.58	3.78	0.81	0.39	0.40

Panel F: Size

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
All	0.11	0.28	(0.11, 0.09)	1.73	1.09	7.98	0.70	0.13	0.16
NYSE	0.07	0.24	(0.10, 0.00)	1.13	-0.45	3.26	0.69	0.04	0.18

Panel G: Trading frictions

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
All	0.17	0.29	(0.03, 0.08)	1.27	0.52	4.14	0.77	0.20	0.15
NYSE	0.12	0.23	(0.02, 0.05)	1.02	0.29	3.65	0.77	0.12	0.11

Panel H: Valuation

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
All	0.31	0.16	(0.00, 0.01)	0.63	0.66	3.74	0.99	0.41	0.56
NYSE	0.24	0.10	(0.00, 0.01)	0.53	1.25	6.34	0.98	0.30	0.49

Table C.7: Impact of decision node: Financials

For each branch of node "financials", we show the mean statistics across sorting variables within each group in separate panels. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: Financing

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.32	0.18	(0.02, 0.07)	1.21	0.67	3.66	0.98	0.74	0.43
Included	0.32	0.17	(0.02, 0.07)	1.19	0.72	3.81	0.98	0.77	0.54

Panel B: Intangibles

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.24	0.14	(0.01, 0.01)	0.79	0.52	4.39	0.84	0.35	0.31
Included	0.22	0.13	(0.00, 0.01)	0.74	0.63	4.44	0.80	0.35	0.38

Panel C: Investment

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.42	0.20	(0.06, 0.10)	1.41	0.71	3.44	1.00	0.96	0.70
Included	0.40	0.19	(0.05, 0.09)	1.36	0.71	3.47	1.00	0.96	0.65

Panel D: Momentum

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.48	0.41	(0.12, 0.17)	2.11	0.59	3.02	0.98	0.79	0.69
Included	0.49	0.38	(0.10, 0.16)	2.07	0.64	3.07	0.98	0.84	0.71

Panel E: Profitability

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.24	0.18	(0.02, 0.04)	0.95	0.68	3.67	0.81	0.41	0.39
Included	0.22	0.17	(0.01, 0.05)	0.90	0.65	3.78	0.82	0.39	0.40

Panel F: Size

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.10	0.25	(0.10, 0.05)	1.47	1.22	10.21	0.70	0.08	0.17
Included	0.09	0.26	(0.11, 0.05)	1.53	0.78	8.66	0.69	0.09	0.16

Panel G: Trading frictions

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.15	0.26	(0.02, 0.07)	1.18	0.60	4.64	0.78	0.16	0.14
Included	0.15	0.26	(0.03, 0.07)	1.17	0.58	4.44	0.76	0.16	0.13

Panel H: Valuation

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.27	0.14	(0.00, 0.01)	0.63	0.94	4.39	0.98	0.31	0.49
Included	0.28	0.13	(0.00, 0.01)	0.60	0.92	4.31	0.99	0.40	0.56

Table C.8: Impact of decision node: Double sort

For each branch of node "double sort", we show the mean statistics across sorting variables within each group in separate panels. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: Financing

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Dependent	0.32	0.17	(0.02, 0.06)	1.18	0.78	3.60	0.99	0.78	0.51
Independent	0.32	0.17	(0.02, 0.06)	1.16	0.73	3.47	0.99	0.77	0.51
Single	0.31	0.21	(0.01, 0.11)	1.25	0.79	3.59	0.94	0.65	0.40

Panel B: Intangibles

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Dependent	0.22	0.12	(0.00, 0.01)	0.72	0.64	4.30	0.81	0.36	0.35
Independent	0.23	0.13	(0.01, 0.01)	0.74	0.69	4.53	0.82	0.35	0.34
Single	0.23	0.16	(0.01, 0.02)	0.83	0.47	4.33	0.84	0.34	0.35

Panel C: Investment

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Dependent	0.41	0.19	(0.06, 0.09)	1.42	0.72	3.40	1.00	0.98	0.66
Independent	0.41	0.19	(0.06, 0.09)	1.38	0.67	3.31	1.00	0.98	0.68
Single	0.41	0.21	(0.04, 0.11)	1.37	0.76	3.52	1.00	0.87	0.68

Panel D: Momentum

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Dependent	0.49	0.40	(0.13, 0.15)	2.12	0.58	2.96	0.98	0.85	0.76
Independent	0.50	0.38	(0.14, 0.16)	2.12	0.64	3.09	0.99	0.86	0.71
Single	0.43	0.35	(0.07, 0.15)	1.90	0.69	3.33	0.97	0.65	0.58

Panel E: Profitability

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Dependent	0.23	0.16	(0.02, 0.04)	0.93	0.71	3.76	0.79	0.43	0.40
Independent	0.24	0.17	(0.02, 0.04)	0.91	0.64	3.71	0.83	0.42	0.39
Single	0.22	0.17	(0.00, 0.05)	0.88	0.78	4.05	0.83	0.31	0.39

Panel F: Size

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Dependent	0.02	0.18	(0.13, 0.08)	2.05	1.75	11.75	0.57	0.10	0.14
Independent	0.09	0.31	(0.06, 0.03)	1.25	0.56	7.85	0.68	0.06	0.15
Single	0.24	0.13	(0.00, 0.04)	1.06	3.51	22.58	0.98	0.10	0.27

Panel G: Trading frictions

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Dependent	0.16	0.25	(0.02, 0.07)	1.22	1.04	5.61	0.81	0.18	0.13
Independent	0.14	0.27	(0.02, 0.06)	1.10	0.57	3.72	0.76	0.12	0.10
Single	0.14	0.24	(0.04, 0.08)	1.18	1.20	7.15	0.72	0.20	0.20

Panel H: Valuation

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Dependent	0.27	0.13	(0.00, 0.00)	0.59	0.81	3.73	0.99	0.37	0.51
Independent	0.27	0.12	(0.00, 0.00)	0.55	0.75	3.62	0.99	0.35	0.53
Single	0.30	0.19	(0.00, 0.03)	0.74	0.88	3.84	0.98	0.35	0.54

Table C.9: Impact of decision node: Breakpoint quantiles (secondary)
For each branch of node "breakpoint quantiles (secondary)", we show the mean statistics across sorting variables within each group in separate panels. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: Financing

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
2	0.32	0.17	(0.01, 0.06)	1.17	0.77	3.58	0.99	0.76	0.51
5	0.33	0.16	(0.02, 0.06)	1.16	0.78	3.46	1.00	0.80	0.50

Panel B: Intangibles

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
2	0.23	0.13	(0.00, 0.01)	0.75	0.68	4.44	0.82	0.36	0.36
5	0.22	0.12	(0.01, 0.01)	0.70	0.64	4.52	0.81	0.35	0.33

Panel C: Investment

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
2	0.41	0.19	(0.06, 0.09)	1.39	0.71	3.40	1.00	0.98	0.69
5	0.41	0.19	(0.06, 0.09)	1.40	0.68	3.29	1.00	0.98	0.65

Panel D: Momentum

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
2	0.49	0.38	(0.12, 0.15)	2.05	0.57	2.94	0.98	0.85	0.72
5	0.51	0.40	(0.15, 0.16)	2.20	0.64	3.03	0.99	0.86	0.75

Panel E: Profitability

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
2	0.24	0.17	(0.02, 0.04)	0.92	0.70	3.75	0.82	0.42	0.41
5	0.23	0.16	(0.02, 0.04)	0.94	0.76	3.86	0.81	0.43	0.39

Panel F: Size

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
2	0.12	0.20	(0.05, 0.05)	1.41	1.93	13.09	0.75	0.09	0.18
5	-0.01	0.33	(0.14, 0.07)	1.52	0.76	7.71	0.50	0.07	0.10

Panel G: Trading frictions

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
2	0.15	0.25	(0.02, 0.07)	1.15	0.70	4.47	0.78	0.16	0.12
5	0.16	0.27	(0.02, 0.07)	1.15	0.59	3.71	0.79	0.14	0.11

Panel H: Valuation

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
2	0.28	0.13	(0.00, 0.00)	0.59	0.77	3.71	0.99	0.39	0.54
5	0.26	0.12	(0.00, 0.00)	0.55	0.75	3.47	0.98	0.33	0.51

Table C.10: Impact of decision node: Rebalancing
For each branch of node "rebalancing", we show the mean statistics across sorting variables within each group in separate panels. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: Financing

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
FF	0.27	0.15	(0.01, 0.06)	1.13	0.74	4.10	0.97	0.69	0.43
monthly	0.33	0.17	(0.04, 0.06)	1.24	0.63	3.80	0.98	0.75	0.53

Panel B: Intangibles

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
FF	0.21	0.13	(0.01, 0.01)	0.78	0.56	4.58	0.81	0.35	0.32
monthly	0.22	0.13	(0.01, 0.01)	0.78	0.61	4.45	0.80	0.37	0.34

Panel C: Investment

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
FF	0.37	0.16	(0.03, 0.06)	1.18	0.70	3.50	1.00	0.95	0.63
monthly	0.45	0.21	(0.09, 0.10)	1.45	0.52	3.19	1.00	0.97	0.71

Panel D: Profitability

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
FF	0.17	0.15	(0.00, 0.02)	0.76	0.59	3.60	0.77	0.28	0.30
monthly	0.18	0.16	(0.01, 0.02)	0.80	0.63	3.88	0.77	0.31	0.32

Panel E: Valuation

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
FF	0.28	0.14	(0.00, 0.01)	0.59	0.83	3.85	0.99	0.37	0.52
monthly	0.29	0.14	(0.00, 0.01)	0.62	0.88	4.23	0.98	0.39	0.55

Table C.11: Impact of decision node: Utilities
For each branch of node "utilities", we show the mean statistics across sorting variables within each group in separate panels. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: Financing

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.32	0.18	(0.02, 0.07)	1.20	0.68	3.69	0.98	0.74	0.48
Included	0.32	0.17	(0.02, 0.07)	1.21	0.72	3.78	0.98	0.76	0.50

Panel B: Intangibles

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.23	0.14	(0.01, 0.01)	0.78	0.62	4.58	0.82	0.34	0.34
Included	0.23	0.14	(0.00, 0.01)	0.78	0.59	4.29	0.82	0.36	0.35

Panel C: Investment

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.42	0.20	(0.06, 0.09)	1.38	0.70	3.44	1.00	0.96	0.65
Included	0.41	0.19	(0.06, 0.09)	1.41	0.72	3.48	1.00	0.96	0.70

Panel D: Momentum

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.50	0.40	(0.11, 0.16)	2.09	0.59	3.07	0.99	0.82	0.71
Included	0.47	0.38	(0.11, 0.16)	2.09	0.62	3.06	0.98	0.81	0.70

Panel E: Profitability

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.23	0.18	(0.02, 0.05)	0.94	0.73	3.87	0.82	0.38	0.37
Included	0.24	0.17	(0.01, 0.04)	0.93	0.72	3.88	0.81	0.42	0.42

Panel F: Size

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.09	0.25	(0.11, 0.05)	1.49	1.05	9.63	0.69	0.08	0.14
Included	0.10	0.26	(0.10, 0.05)	1.50	0.98	9.41	0.70	0.09	0.19

Panel G: Trading frictions

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.17	0.27	(0.03, 0.07)	1.21	0.55	4.70	0.80	0.18	0.14
Included	0.13	0.25	(0.02, 0.06)	1.14	0.65	4.69	0.75	0.14	0.12

Panel H: Valuation

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
Excluded	0.28	0.14	(0.00, 0.01)	0.61	0.90	4.37	0.99	0.37	0.56
Included	0.27	0.14	(0.00, 0.01)	0.63	0.96	4.46	0.98	0.35	0.49

Table C.12: Impact of decision node: Stock age restriction
For each branch of node "stock age restriction", we show the mean statistics across sorting variables within each group in separate panels. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: Financing

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.31	0.18	(0.02, 0.07)	1.25	0.64	3.62	0.98	0.73	0.47
2	0.29	0.16	(0.01, 0.08)	1.22	0.77	4.03	0.98	0.71	0.50

Panel B: Intangibles

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.16	0.12	(0.00, 0.01)	0.75	0.56	4.68	0.76	0.32	0.27
2	0.15	0.11	(0.00, 0.01)	0.75	0.69	4.96	0.79	0.25	0.26

Panel C: Investment

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.43	0.20	(0.06, 0.09)	1.39	0.67	3.34	1.00	0.96	0.67
2	0.39	0.18	(0.05, 0.09)	1.34	0.69	3.42	1.00	0.96	0.68

Panel D: Momentum

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.51	0.42	(0.12, 0.19)	2.28	0.70	3.14	0.99	0.83	0.72
2	0.47	0.41	(0.11, 0.17)	2.17	0.67	3.16	0.97	0.81	0.70

Panel E: Profitability

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.24	0.19	(0.02, 0.04)	0.94	0.62	3.61	0.82	0.40	0.38
2	0.23	0.17	(0.01, 0.04)	0.91	0.74	3.91	0.81	0.40	0.41

Panel F: Size

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.09	0.24	(0.09, 0.05)	1.42	1.13	10.53	0.70	0.07	0.12
2	0.10	0.28	(0.12, 0.05)	1.57	0.92	8.68	0.70	0.10	0.21

Panel G: Trading frictions

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.19	0.31	(0.03, 0.09)	1.44	0.77	4.86	0.79	0.22	0.17
2	0.16	0.30	(0.03, 0.09)	1.43	0.77	5.12	0.77	0.19	0.17

Panel H: Valuation

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.29	0.15	(0.00, 0.01)	0.64	0.79	3.81	0.98	0.40	0.54
2	0.27	0.13	(0.00, 0.01)	0.57	0.90	4.27	0.99	0.36	0.53

Table C.13: Impact of decision node: Price restriction

For each branch of node "price restriction", we show the mean statistics across sorting variables within each group in separate panels. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: Financing

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.33	0.18	(0.02, 0.08)	1.28	0.79	3.79	0.98	0.76	0.47
1	0.33	0.18	(0.02, 0.07)	1.22	0.64	3.50	0.98	0.76	0.48
5	0.31	0.16	(0.01, 0.04)	1.10	0.48	3.30	0.98	0.75	0.51

Panel B: Intangibles

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.23	0.14	(0.01, 0.02)	0.82	0.65	4.44	0.82	0.35	0.35
1	0.23	0.14	(0.01, 0.01)	0.79	0.56	4.08	0.82	0.36	0.35
5	0.22	0.13	(0.00, 0.01)	0.70	0.31	3.74	0.82	0.34	0.34

Panel C: Investment

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.42	0.20	(0.06, 0.11)	1.46	0.77	3.62	1.00	0.96	0.67
1	0.42	0.20	(0.06, 0.10)	1.40	0.65	3.21	1.00	0.96	0.68
5	0.40	0.18	(0.05, 0.08)	1.31	0.62	3.13	1.00	0.96	0.66

Panel D: Momentum

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.48	0.40	(0.12, 0.16)	2.15	0.60	3.20	0.98	0.81	0.70
1	0.49	0.40	(0.12, 0.17)	2.13	0.64	2.99	0.98	0.82	0.71
5	0.47	0.38	(0.10, 0.15)	1.98	0.52	2.67	0.99	0.82	0.70

Panel E: Profitability

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.24	0.18	(0.02, 0.05)	0.97	0.71	3.85	0.82	0.40	0.40
1	0.24	0.18	(0.02, 0.05)	0.94	0.68	3.59	0.82	0.40	0.40
5	0.23	0.17	(0.02, 0.04)	0.90	0.60	3.49	0.80	0.40	0.39

Panel F: Size

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.13	0.28	(0.11, 0.08)	1.75	1.45	9.26	0.71	0.13	0.18
1	0.10	0.27	(0.11, 0.05)	1.39	0.22	5.09	0.71	0.10	0.18
5	0.04	0.23	(0.11, 0.00)	1.17	-0.86	3.79	0.66	0.03	0.13

Panel G: Trading frictions

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.15	0.26	(0.03, 0.07)	1.25	0.73	4.94	0.76	0.17	0.14
1	0.14	0.26	(0.03, 0.07)	1.17	0.53	3.82	0.76	0.16	0.13
5	0.15	0.25	(0.02, 0.06)	1.10	0.36	3.53	0.80	0.16	0.13

Panel H: Valuation

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
0	0.28	0.15	(0.00, 0.01)	0.67	0.98	4.46	0.99	0.37	0.54
1	0.28	0.14	(0.00, 0.01)	0.62	0.78	3.68	0.99	0.37	0.53
5	0.26	0.12	(0.00, 0.00)	0.54	0.58	3.29	0.98	0.34	0.50

Table C.14: Impact of decision node: Positive book equity

For each branch of node "positive book equity", we show the mean statistics across sorting variables within each group in separate panels. Each panel contains the mean (Mean, in %), skewness (Skew.), and kurtosis (Kurt.) of the premia. Furthermore, they contain the non-standard error (NSE, in %) and the relative number of significant deviations to the left and right of the median using a 5% significance level (Left-right). The table also shows the ratio of the dispersion of premia relative to the average time-series standard error (Ratio). Columns Pos. and Sig. show the relative number of positive premia and t-statistics larger than 1.96. The last column (Mon.) shows the relative number of monotonically increasing portfolio sorts following Patton and Timmermann (2010) and testing all possible pairs at a 10% significance level.

Panel A: Financing

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.32	0.17	(0.02, 0.07)	1.20	0.69	3.72	0.98	0.75	0.49
Yes	0.32	0.17	(0.02, 0.07)	1.21	0.69	3.74	0.98	0.75	0.49

Panel B: Intangibles

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.23	0.14	(0.01, 0.01)	0.78	0.61	4.46	0.82	0.35	0.34
Yes	0.23	0.14	(0.01, 0.01)	0.78	0.60	4.45	0.82	0.35	0.35

Panel C: Investment

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.41	0.19	(0.06, 0.09)	1.40	0.71	3.46	1.00	0.96	0.67
Yes	0.41	0.19	(0.06, 0.09)	1.40	0.71	3.46	1.00	0.96	0.67

Panel D: Momentum

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.48	0.39	(0.11, 0.16)	2.09	0.61	3.07	0.98	0.81	0.70
Yes	0.48	0.39	(0.11, 0.16)	2.10	0.61	3.07	0.98	0.81	0.70

Panel E: Profitability

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.23	0.18	(0.02, 0.05)	0.94	0.69	3.79	0.81	0.40	0.39
Yes	0.23	0.18	(0.02, 0.05)	0.94	0.69	3.81	0.82	0.40	0.40

Panel F: Size

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.09	0.26	(0.11, 0.05)	1.50	1.00	9.48	0.70	0.08	0.17
Yes	0.09	0.26	(0.11, 0.05)	1.50	1.03	9.54	0.70	0.08	0.17

Panel G: Trading frictions

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.15	0.26	(0.03, 0.07)	1.18	0.62	4.60	0.77	0.16	0.13
Yes	0.15	0.26	(0.03, 0.07)	1.18	0.62	4.60	0.77	0.16	0.13

Panel H: Valuation

Branch	Mean	NSE	Left-right	Ratio	Skew.	Kurt.	Pos.	Sig.	Mon.
No	0.28	0.14	(0.00, 0.01)	0.62	0.92	4.38	0.99	0.36	0.53
Yes	0.27	0.14	(0.00, 0.01)	0.62	0.91	4.37	0.98	0.36	0.52

D FF5 and Q5 alphas and individual decision nodes

In Tables D.1 to D.3, we present the impact of decision nodes on CAPM-adjusted, Fama and French (2015)-adjusted (FF5), and Hou et al. (2021)-adjusted (Q5) premia.

Table D.1: CAPM-adjusted returns: Mean absolute differences and correlations.
This table shows mean absolute differences (Panel A, in %) and correlations (Panel B) of the CAPM-adjusted time series of premia across decision nodes. For each decision node, we compare time-series pairs that differ only in the specific node. Then, we take the mean for each node-sorting variable combination. The two panels show means for over all categories and individual categories separately. By construction, some entries do not produce variation and are left empty.

Panel A: Mean absolute differences

Node	All	Fin.	Int.	Inv.	Mom.	Pro.	Siz.	Tra.	Val.
BP: Quantiles (main)	1.06	0.90	1.11	0.95	1.03	1.07	1.39	1.30	1.01
Weighting scheme	0.99	0.96	1.04	0.95	1.00	1.00	0.63	0.92	1.06
Positive earnings	0.97	0.85	0.99	0.83	0.79	1.27	1.38	1.08	0.87
Size restriction	0.85	0.68	0.85	0.69	0.78	0.89	1.91	1.15	0.79
Sorting variable lag	0.84	0.56	0.53	0.63	1.68	0.64	1.84	1.52	0.57
BP: Exchanges	0.84	0.69	0.87	0.67	0.67	0.90	1.58	1.14	0.78
Financials	0.75	0.46	0.75	0.61	0.60	1.05	0.79	0.76	0.86
Double sort	0.70	0.42	0.67	0.45	0.51	0.74	2.90	1.32	0.52
BP: Quantiles (second)	0.69	0.53	0.69	0.53	0.59	0.71	1.84	1.05	0.60
Rebalancing	0.59	0.59	0.59	0.62		0.59			0.55
Utilities	0.48	0.37	0.36	0.33	0.42	0.66	0.60	0.64	0.54
Stock-age restriction	0.43	0.46	0.39	0.43	0.32	0.47	0.65	0.52	0.35
Price restriction	0.37	0.31	0.39	0.31	0.32	0.40	0.77	0.48	0.35
Positive book equity	0.08	0.07	0.08	0.07	0.07	0.08	0.13	0.10	0.07

Panel B: Correlations

Node	All	Fin.	Int.	Inv.	Mom.	Pro.	Siz.	Tra.	Val.
Weighting scheme	0.86	0.80	0.84	0.81	0.88	0.88	0.97	0.91	0.90
Positive earnings	0.87	0.85	0.88	0.83	0.92	0.80	0.83	0.91	0.93
Sorting variable lag	0.88	0.92	0.95	0.90	0.71	0.95	0.64	0.74	0.96
BP: Quantiles (main)	0.90	0.87	0.88	0.86	0.92	0.92	0.87	0.90	0.94
Size restriction	0.91	0.90	0.91	0.90	0.93	0.92	0.75	0.90	0.95
Financials	0.92	0.95	0.91	0.91	0.96	0.85	0.93	0.95	0.93
BP: Exchanges	0.92	0.91	0.92	0.91	0.95	0.92	0.81	0.91	0.95
Rebalancing	0.92	0.89	0.93	0.89		0.95			0.96
BP: Quantiles (second)	0.94	0.93	0.94	0.94	0.95	0.95	0.72	0.91	0.97
Double sort	0.94	0.96	0.94	0.95	0.97	0.94	0.63	0.87	0.97
Utilities	0.97	0.97	0.98	0.98	0.98	0.94	0.95	0.97	0.97
Stock-age restriction	0.97	0.95	0.97	0.96	0.99	0.97	0.94	0.97	0.99
Price restriction	0.97	0.97	0.97	0.97	0.98	0.97	0.93	0.97	0.98
Positive book equity	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00

Table D.2: FF5-adjusted returns: Mean absolute differences and correlations.
This table shows mean absolute differences (Panel A, in %) and correlations (Panel B) of the Fama and French (2015)-adjusted time series of premia across decision nodes. For each decision node, we compare time-series pairs that differ only in the specific node. Then, we take the mean for each node-sorting variable combination. The two panels show means for over all categories and individual categories separately. By construction, some entries do not produce variation and are left empty.

Panel A: Mean absolute differences

Node	All	Fin.	Int.	Inv.	Mom.	Pro.	Siz.	Tra.	Val.
BP: Quantiles (main)	1.02	0.87	1.07	0.93	1.01	1.02	1.35	1.25	0.93
Weighting scheme	0.96	0.94	1.02	0.94	1.00	0.96	0.58	0.89	1.01
Positive earnings	0.89	0.77	0.93	0.78	0.77	1.07	1.35	1.01	0.82
Sorting variable lag	0.84	0.56	0.53	0.63	1.66	0.64	1.87	1.52	0.57
Size restriction	0.83	0.66	0.84	0.67	0.76	0.87	1.91	1.13	0.77
BP: Exchanges	0.81	0.67	0.84	0.66	0.66	0.85	1.58	1.12	0.73
Financials	0.71	0.45	0.70	0.56	0.60	0.96	0.78	0.75	0.83
Double sort	0.68	0.42	0.66	0.45	0.51	0.73	2.21	1.29	0.52
BP: Quantiles (second)	0.66	0.52	0.66	0.52	0.59	0.66	1.72	0.99	0.58
Rebalancing	0.60	0.60	0.59	0.63		0.61			0.56
Utilities	0.47	0.36	0.36	0.33	0.41	0.63	0.61	0.63	0.53
Stock-age restriction	0.42	0.43	0.39	0.41	0.32	0.46	0.66	0.52	0.34
Price restriction	0.37	0.31	0.38	0.31	0.32	0.40	0.77	0.47	0.34
Positive book equity	0.08	0.08	0.08	0.07	0.07	0.08	0.14	0.10	0.07

Panel B: Correlations

Node	All	Fin.	Int.	Inv.	Mom.	Pro.	Siz.	Tra.	Val.
Weighting scheme	0.79	0.72	0.79	0.75	0.86	0.81	0.92	0.88	0.76
Positive earnings	0.83	0.80	0.84	0.81	0.92	0.78	0.72	0.88	0.84
Sorting variable lag	0.84	0.88	0.93	0.86	0.68	0.91	0.39	0.68	0.91
BP: Quantiles (main)	0.85	0.82	0.85	0.82	0.91	0.87	0.79	0.88	0.86
Size restriction	0.86	0.85	0.87	0.86	0.92	0.86	0.49	0.85	0.86
Rebalancing	0.88	0.84	0.90	0.84		0.91			0.89
BP: Exchanges	0.88	0.87	0.88	0.88	0.94	0.88	0.68	0.88	0.89
Financials	0.89	0.93	0.90	0.91	0.95	0.80	0.88	0.94	0.83
Double sort	0.91	0.94	0.91	0.94	0.96	0.91	0.44	0.84	0.93
BP: Quantiles (second)	0.91	0.90	0.91	0.91	0.95	0.92	0.63	0.90	0.92
Utilities	0.95	0.96	0.97	0.97	0.98	0.92	0.92	0.96	0.93
Stock-age restriction	0.96	0.93	0.97	0.94	0.98	0.96	0.90	0.96	0.97
Price restriction	0.96	0.95	0.96	0.96	0.98	0.96	0.88	0.96	0.96
Positive book equity	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00

Table D.3: Q5-adjusted returns: Mean absolute differences and correlations.
This table shows mean absolute differences (Panel A, in %) and correlations (Panel B) of the Hou et al. (2021)-adjusted time series of premia across decision nodes. For each decision node, we compare time-series pairs that differ only in the specific node. Then, we take the mean for each node-sorting variable combination. The two panels show means for over all categories and individual categories separately. By construction, some entries do not produce variation and are left empty.

Panel A: Mean absolute differences

Node	All	Fin.	Int.	Inv.	Mom.	Pro.	Siz.	Tra.	Val.
BP: Quantiles (main)	1.03	0.88	1.08	0.93	0.98	1.04	1.36	1.26	0.97
Weighting scheme	0.97	0.94	1.02	0.93	1.00	0.97	0.57	0.89	1.03
Positive earnings	0.94	0.82	0.96	0.80	0.78	1.20	1.37	1.06	0.86
Size restriction	0.84	0.66	0.84	0.68	0.77	0.88	1.91	1.13	0.78
Sorting variable lag	0.83	0.56	0.53	0.62	1.65	0.63	1.86	1.52	0.57
BP: Exchanges	0.82	0.68	0.85	0.66	0.66	0.88	1.59	1.13	0.75
Financials	0.73	0.45	0.73	0.59	0.60	1.02	0.79	0.75	0.84
Double sort	0.68	0.42	0.66	0.45	0.51	0.73	2.23	1.29	0.52
BP: Quantiles (second)	0.66	0.52	0.67	0.52	0.59	0.67	1.72	0.99	0.58
Rebalancing	0.60	0.60	0.60	0.63		0.61			0.56
Utilities	0.47	0.36	0.36	0.33	0.41	0.64	0.61	0.63	0.53
Stock-age restriction	0.42	0.45	0.39	0.43	0.33	0.47	0.66	0.53	0.34
Price restriction	0.37	0.31	0.39	0.31	0.32	0.40	0.77	0.48	0.35
Positive book equity	0.08	0.08	0.08	0.07	0.07	0.08	0.14	0.10	0.07

Panel B: Correlations

Node	All	Fin.	Int.	Inv.	Mom.	Pro.	Siz.	Tra.	Val.
Weighting scheme	0.82	0.74	0.81	0.75	0.84	0.83	0.94	0.89	0.86
Positive earnings	0.84	0.80	0.85	0.80	0.90	0.74	0.71	0.88	0.91
Sorting variable lag	0.85	0.89	0.94	0.86	0.62	0.91	0.41	0.69	0.95
BP: Quantiles (main)	0.87	0.83	0.86	0.82	0.89	0.88	0.80	0.88	0.92
Size restriction	0.88	0.86	0.89	0.86	0.91	0.88	0.53	0.86	0.92
BP: Exchanges	0.90	0.88	0.90	0.88	0.93	0.89	0.69	0.88	0.94
Financials	0.90	0.94	0.90	0.89	0.94	0.82	0.88	0.94	0.91
Rebalancing	0.90	0.85	0.91	0.85		0.93			0.95
Double sort	0.92	0.95	0.93	0.94	0.96	0.92	0.46	0.85	0.96
BP: Quantiles (second)	0.92	0.91	0.93	0.91	0.94	0.93	0.64	0.91	0.95
Utilities	0.96	0.96	0.97	0.97	0.97	0.93	0.92	0.96	0.96
Stock-age restriction	0.96	0.93	0.97	0.94	0.98	0.96	0.91	0.97	0.99
Price restriction	0.96	0.96	0.96	0.96	0.97	0.96	0.88	0.96	0.98
Positive book equity	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00