

Essays on Empirical Asset Pricing

Jensen, Theis Ingerslev

Document Version

Final published version

DOI:

[10.22439/phd.13.2023](https://doi.org/10.22439/phd.13.2023)

Publication date:

2023

License

Unspecified

Citation for published version (APA):

Jensen, T. I. (2023). *Essays on Empirical Asset Pricing*. Copenhagen Business School [Phd]. PhD Series No. 13.2023 <https://doi.org/10.22439/phd.13.2023>

[Link to publication in CBS Research Portal](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 30. Mar. 2025



COPENHAGEN BUSINESS SCHOOL
SOLBJERG PLADS 3
DK-2000 FREDERIKSBERG
DANMARK

WWW.CBS.DK

ISSN 0906-6934

Print ISBN: 978-87-7568-171-6
Online ISBN: 978-87-7568-172-3

DOI: <https://doi.org/10.22439/phd.13.2023>

ESSAYS ON EMPIRICAL ASSET PRICING

Theis Ingerslev Jensen
**ESSAYS ON EMPIRICAL ASSET
PRICING**

Finance

PhD Series 13.2023

CBS  COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

PhD Series 13.2023

Essays on Empirical Asset Pricing

Theis Ingerslev Jensen

A thesis presented for the degree of Doctor of Philosophy

Supervisor: Lasse Heje Pedersen

Ph.D. School in Economics and Management
Copenhagen Business School

Theis Ingerslev Jensen
Essays on Empirical Asset Pricing

First edition 2023
Ph.D. Series 13.2023

© Theis Ingerslev Jensen

ISSN 0906-6934

Print ISBN: 978-87-7568-171-6
Online ISBN: 978-87-7568-172-3

DOI: <https://doi.org/10.22439/phd.13.2023>

All rights reserved.

Copies of text contained herein may only be made by institutions that have an agreement with COPY-DAN and then only within the limits of that agreement. The only exception to this rule is short excerpts used for the purpose of book reviews.

Abstract

This thesis investigates the relationship between risk and return in the cross-section of stocks. The thesis consists of three papers that can be read independently. These papers are the result of my Ph.D. studies at the Department of Finance, the Center of Financial Frictions (FRIC), and the Center of Big Data in Finance (BIGFI) at CBS.

The first paper, Subjective Risk and Return, investigate the subjective risk-return tradeoff. To do so, I introduce a novel data set of subjective risk and return expectations at the individual stock level. I find that the required compensation for risk is high, but nevertheless, the realized compensation for risk is low. I show that this difference arises because cash flow expectations are systematically too high for risky stocks, which can be explained by investors suffering from optimism bias. As a result of the low realized compensation for risk, I show that risk cannot explain the realized return of most equity factors, and that the best asset pricing models for explaining realized returns are the worst ones for explaining subjective risk compensation.

The second paper, Is There a Replication Crisis in Finance (co-authored with Bryan Kelly and Lasse Heje Pedersen), tests whether equity factor research in finance is robust to scientific replication. In particular, we build a global data set of stock returns and characteristics to replicate 153 equity factors in 93 different countries. Further, we develop and estimate a Bayesian model of factor replication, which can handle the multiple testing of many hypotheses, and the issue of publication bias. Our main result is that most equity factors can be replicated. Further, we show that most factors work well out-of-sample, that is, in time periods and countries different from the ones studied in the original paper. We also show that the 153 factors can be grouped into 13 themes, most of which matter for the tangency portfolio.

The third paper, Machine Learning and the Implementable Efficient Frontier (co-authored with Bryan Kelly, Semyon Malamud, and Lasse Heje Pedersen), develops a framework that integrates trading-cost-aware portfolio optimization with machine learning (ML). We show theoretically how to solve the optimal portfolio problem for an investor that faces trading costs when returns are predictable by a general function of security characteristics. In addition, we show to implement this solution via a machine learning methodology that learns directly about portfolio weight (rather than returns). Empiri-

cally, we find that our method leads to significant out-of-sample gains relative to various sophisticated benchmarks. Finally, our method gives a novel view of which security characteristics are economically important.

Acknowledgements

Completing my Ph.D. has been a tremendously rewarding experience, but it would not have been possible without the help of several people, only some of whom I can highlight here.

First and foremost, I am incredibly grateful for the help and support from my advisor, Lasse Heje Pedersen. Lasse has dedicated more time to my development than anyone can reasonably expect from an advisor. Furthermore, by inviting me to be his co-author, he has shown me firsthand what it takes to be a world-class researcher. He has taught me to be critical, curious, and always search for the truth. In addition, Lasse has this remarkable ability to push you research-wise while simultaneously reminding you that there is a life outside of research. Working with Lasse has been an honor and a privilege.

Second, I want to give a special thanks to Bryan Kelly, who sponsored my exchange stay at Yale. Bryan immediately took me under his wings, and he has been advising me ever since. Three years after I first walked into his office, I am proud to say that he is my co-author, future colleague, and friend.

Third, I want to thank my colleagues at Copenhagen Business School. In particular, David Lando, who first hired me as a research assistant and then as a Ph.D. student, has been instrumental in creating a good environment at the department. I have also enjoyed being part of the fantastic group of Ph.D. students at CBS. My two office mates, Julian Terstegge and Alessandro Spina, deserve special thanks for enduring me for these past years and for countless interesting discussions.

Last, but definitely not least, I am forever grateful to my friends and family. My friends ensured that my life was never dull, and my family ensured that I always had unconditional support. Still, I owe the greatest debt to my fiancé, Othilie, who has picked me up when I was down, cheered me on when I was up, and always reminded what's important in life. Words cannot describe how much you mean to me.

Summaries

1 Summaries in English

Subjective Risk and Return

Risk-averse investors require a higher return to invest in riskier stocks. Therefore, the required return of a stock depends on its risk, as perceived by investors, and the compensation investors require for taking risk. The standard approach is to infer required returns from an asset pricing model such as the Capital Asset Pricing Model (CAPM) of Sharpe (1964), Lintner (1965), and Mossin (1966) or the 3-factor model (FF3) of Fama and French (1993a). According to the CAPM, the required return on a stock depends on its market beta, while FF3 predicts that a stock's required return depends on its exposure to market risk, while FF3 predicts that it also depends on its exposure to size and value risk. However, we have many different asset pricing models, and it is unclear which one best describes the behavior of actual investors.

Instead of relying on an asset pricing model, I infer required returns from observable subjective risk and return expectations. The subjective risk ratings give me a direct proxy for which stocks investors perceive as risky but not the risk premium required by investors. To estimate this subjective risk premium, I regress subjective expected returns on subjective risk. My final estimate of a stock's required return is then the product of its subjective risk and the subjective risk premium.

My first main result is that the subjective risk premium is high. In particular, I estimate that investors require a 12% higher return each year for investing in the riskiest stocks relative to the safest. Nevertheless, the realized risk premium is low in the sense that the realized return is only 3% higher for the riskiest stocks relative to the safest. This discrepancy arises because cash flow forecasts are systematically too high for the riskiest stocks, which I show is consistent with investors suffering from optimism bias. I make two additional findings based on the required return estimates: First, most equity factors have a negative required return despite a positive realized return, which is inconsistent with equity factors being driven by rational compensation for risk. Second, I show that recent asset pricing models, such as the 5-factor model of Fama and French (2015), explain

realized returns well but required returns poorly. In contrast, traditional models, such as the CAPM, explain required returns well but realized returns poorly.

Is There a Replication Crisis in Finance

Several research fields, such as medicine (Ioannidis, 2005) and psychology (Nosek et al., 2012), face issues that published research findings cannot be reproduced, a so-called replication crisis. These concerns have also been raised in financial economics and are specifically directed toward the literature that uses equity factors to study return predictability. For example, Hou et al. (2020a) argues that most equity factors lack internal validity because they cannot be replicated after small modifications to the factor construction. Further, Harvey et al. (2016a) argues that most equity factors lack external validity because they result from “*p*-hacking” whereby researchers try multiple hypotheses but only select the best-performing ones for publication. We examine both of these challenges and argue that neither is tenable.

To examine these challenges, we build a large new data set of global stock returns and firm characteristics, which we use to replicate 153 equity factors in 93 countries. The majority of papers that proposes a new equity factor only test the factor on US data, so our global data set provides a useful out-of-sample test to evaluate the robustness of the original result. In addition, we develop a Bayesian model of factor replication, which estimates the expected performance of all factors jointly, thereby utilizing information about the performance of related factors. By modeling all factors jointly, the Bayesian model naturally accounts for issues with multiple testing. Furthermore, we show how the model can account for publication bias even without observing all factors that have been tested.

We find that US equity factors have a high degree of internal validity in the sense that 82.4% of the factors remain significant after modification to the factor constructions, which makes all factors consistent and avoids putting too much weight on small stocks, thereby making the factor more implementable. This conclusion is robust to accounting for multiple testing via our Bayesian model. We also show that factors show a high degree of external validity in the sense that they replicate well in countries and time periods that differ from the ones studied in the original paper. Finally, we show that the 153 factors can be clustered into 13 distinct themes, most of which are a significant part of the tangency portfolio.

Machine Learning and the Implementable Efficient Frontier

There is a growing literature that uses machine learning (ML) techniques to predict stock returns (e.g., Gu et al. (2020a)). These models are typically designed to predict returns

without considering trading cost, and, as a result, they tend to focus on short-lived characteristics that work well for small and illiquid stocks (see, e.g., Avramov et al. (2021)). However, real-world investors face trading cost, so the ML-implied portfolios are not implementable in practice. We propose that investment strategies should be evaluated based on their expected returns out-of-sample after trading cost for different levels of risk, which we call the “implementable efficient frontier.”

We develop a framework that integrates trading-cost-aware portfolio optimization with ML. Our key theoretical insight is that the optimal portfolio with trading cost is a compromise between the portfolio inherited from the last period and a so-called “Aim Portfolio,” which is the weighted average of the optimal portfolios without trading costs over all future periods. Further, we show that the optimal weight to put on a stock depends not only on its expected return over the next period (as would be the case without trading costs) but on its expected return over all future periods. Furthermore, how aggressively the investor should trade toward the desired portfolio weight depends on the stock’s liquidity and the investor’s wealth.

We consider two different approaches of using ML to implement the theoretical results. The first approach, Multiperiod-ML, uses ML to predict returns over multiple future horizons and uses the theoretical result to turn these predictions into portfolio weights. The second approach, Portfolio-ML, uses ML to learn directly about a stock’s weight in the Aim Portfolio as a function of its characteristics. Empirically, we compare our two methods to several existing benchmark and shows that the out-of-sample gains for a mean-variance investor can be substantial. In particular, we show that Portfolio-ML outperforms the second-best alternative by 20% in net Sharpe ratio and 60% in utility terms. In addition, and consistent with the theoretical analysis, we show that slow-moving features of a stock, such as its value and quality characteristics, are the most important determinants of its weight in the aim portfolio.

2 Summaries in Danish

Subjective Risk and Return

Risiko-averse investorer kræver et højere afkast for mere risikable aktier. Derfor afhænger det krævede afkast på en aktie af hvor risikable investorer syntes aktien er og hvor meget compensation de kræver for at tage risiko. Den klassiske metode til at estimere det krævede afkast på en aktie er at bruge en asset pricing model såsom Capital Asset Pricing Modellen (CAPM) fra Sharpe (1964), Lintner (1965), and Mossin (1966) eller 3-faktor modellen (FF3) fra Fama and French (1993a). Ifølge CAPM afhænger det krævede afkast på en aktie af dens markedseksponeering, hvorimod FF3 siger at det også afhænger af

aktien size og value eksponering. Problemet er at vi har mange forskellige modeller og det er uklart hvilken model, der bedst beskriver faktiske investorer.

I stedet for at bruge en asset pricing model finder jeg det krævede afkast ved hjælp af observerbare subjektive risiko og afkast forventninger. De subjektive risiko vurderinger giver mig en direkte proxy for hvilke aktier investorer anser som risikable, men jeg mangler stadig den risiko præmie som investorer kræver. Jeg estimerer denne subjektive risiko-præmie ved at regressere de subjektive afkastforventninger på den subjektive risiko. Mit endelige estimat for det krævede afkast på en aktie er så produktet af dens subjektive risiko og den subjektive risikopræmie.

Mit første hovedresultat er at den subjektive risikopræmie er høj. Specifikt estimerer jeg at investorer kræver et 12% højere afkast hvert år for at investere i de mest risikable aktier relativt til de sikreste. Ikke desto mindre er den realiserede risikopræmie lav i den forstand at det realiserede afkast kun er 3% højere for de mest risikable aktier relativt til de sikreste. Denne forskel opstår fordi at forventningerne til indtjeningen for de mest risikable aktier er systematisk for høje, hvilket jeg viser er konsistent med investorer, der lider af en optimisme bias. Jeg finder to yderligere resultater baseret mine nye mål for det krævede afkast på en aktie: For det første viser jeg at de fleste aktiefaktorer har et negativt krævet afkast på trods af et positivt realiseret afkast, hvilket er inkonsistent med en forklaring om at disse faktorer er drevet af en rationel compensation for risiko. For det andet viser jeg at nyere asset pricing modeller, såsom 5-faktor modellen fra Fama and French (2015), forklarer realiserede afkast godt, men krævede afkast dårligt. Omvendt forklarer traditionelle modeller, såsom CAPM, krævede afkast godt, men realiserede afkast dårligt.

Is There a Replication Crisis in Finance

Mange forskningsfelter, såsom medicin (Ioannidis, 2005) og psykologi (Nosek et al., 2012), har problemer med at publicerede forskningsresultater ikke kan blive genskabt, en såkaldt replikationskrise. Denne bekymring er også blevet rejst i finansiel økonomi og er specifikt blevet rettet mod litteraturen, der bruger aktiefaktorer til at undersøge om man kan prædiktere aktieafkast. For eksempel argumentere Hou et al. (2020a) for at de fleste aktiefaktorer mangler intern validitet fordi de ikke kan blive replikeret efter små ændringer i forhold til hvordan de bliver lavet. Ydermere argumentere Harvey et al. (2016a) for at de fleste aktiefaktorer mangler ekstern validitet fordi de er resultatet af “*p*-hacking” hvormed forskere tester forskellige hypoteser men kun vælger at publicere de bedste. Vi undersøger begge disse kritikpunkter og finder at ingen af dem er holdbare.

For at undersøge disse kritikpunkter bygger vi et stort nyt datasæt med globale aktie afkast og firma karakteristikker, som vi bruger til at replikere 153 aktiefaktorer i 93 lande.

De fleste forskningsartikler, der studere aktiefaktorer tester kun faktoren på amerikansk data, så vores globale datasæt giver et nyttig “out-of-sample” test til at vurdere robustheden af det oprindelige resultat. Derudover udvikler vi en Bayesiansk model for faktorreplikation, som estimerer den forventede afkast på alle aktier simultant, hvorved den bruger information for afkastet på tæt relaterede faktorer. Ved at modellere alle faktorer på samme tid, kan den Bayesianske model tage højde for at vi tester flere hypoteser på samme tid. Derudover viser vi hvordan denne model kan bruges til at tage højde for publikations bias selv uden at observere alle faktorer, der er blevet testet.

Vi finder at amerikanske aktiefaktorer udviser en høj grad af intern validitet på den måde at 82.4% af de aktiefaktorer vi tester forbliver signifikante efter vi ændre deres konstruktion ved at gøre dem mere konsistente og undgå at putte for høj vægt på små aktier, hvilket gør at faktorerne er lettere at implementere. Denne konklusion er robust når man tager højde for at vi har testet mange faktorer på samme tid gennem vores Bayesianske model. Vi viser også at faktorer udviser en høj grad af ekstern validitet på den måde at de replikere godt i lande og tidsperioder, der afviger fra dem, som blev studeret i den originale artikel. Til sidst viser vi at de 153 faktorer kan blive grupperet ind i 13 forskellige temaer, hvoraf de fleste er en signifikant del af tangent porteføljen.

Machine Learning and the Implementable Efficient Frontier

Der er en voksende litteratur, som bruger maskinlæringsteknikker (ML) til at forudsige aktieafkast (f.eks. Gu et al. (2020a)). Disse modeller er typisk designet til at forudsige afkast uden at tage højde for handelsomkostninger, hvilket gør at de har en tendens til at fokusere på kortsigtet profitabilitet, der fungerer godt for små og illikvide aktier (se f.eks. Avramov et al. (2021)). Investorer i den virkelige verden bliver dog nødt til at forholde sig til handelsomkostninger og derfor er disse ML porteføljer ikke mulige at implementere i praksis. Vi foreslår i stedet at investeringsstrategier skal vurderes på deres forventede afkast out-of-sample efter handelsomkostninger for forskellige risikoniveauer, hvilket vi kalder den “implementable efficient frontier.”

Vi udvikler et framework, der integrerer handelsomkostningsbevidst porteføljeoptimering med ML. Vores centrale teoretiske indsigt er at den optimale portefølje med handelsomkostninger er et kompromis mellem den portefølje investoren arvede fra den tidligere periode og en såkaldt “mål portefølje,” som er det vægtede gennemsnit af de optimale porteføljer uden handelsomkostninger i alle fremtidige perioder. Derudover viser vi at den optimale vægtning af en aktie ikke kun afhænger af den forventede afkast over den næste periode (hvilket er tilfældet uden handelsomkostninger), men af dens forventede afkast over alle fremtidige perioder. Desuden viser vi at farten, der skal handles mod den ønskede vægt afhænger af aktiens likviditet samt investorens formue.

Vi analyserer to forskellige måder at bruge ML til at implementere vores teoretiske resultater. Den første tilgang, Multiperiod-ML, bruger ML til at forudsige aktieafkast over flere fremtidige horisonter og bruger det teoretiske resultat til at omdanne disse forudsigelser til porteføljevægte. Den anden tilgang, Portfolio-ML, bruger ML til at lære direkte om en akties vægt i mål porteføljen som en funktion af aktiens karakteristika. Empirisk sammenligner vi vores to metoder med flere eksisterende benchmarks og viser at gevinsterne for en investor med en mean-variance nyttefunktion kan være betydelige. Specielt Portfolio-ML klarer sig godt og slår det næstbedste alternativ med 20% i net Sharpe ratio og 60% i forhold til nytteværdien. Derudover viser vi, i overensstemmelse med den teoretiske analyse, at aktie karakteristika, der ændrer sig langsomt, såsom en aktie value og quality karakteristika, er de vigtigste til at bestemme, hvor stor en vægt aktien får i mål porteføljen.

Contents

Abstract	iii
Acknowledgements	v
Summaries	vii
1 Summaries in English	vii
2 Summaries in Danish	ix
1 Subjective Risk and Return	1
1 Theory	7
1.1 Setup	7
1.2 Results	8
2 Data	13
3 The Subjective and Realized Risk Premium	17
3.1 The Subjective Risk Premium is High	18
3.2 The Realized Risk Premium is Low	20
4 Subjective Risk and Cash Flow Optimism	22
4.1 Subjective Risk Predicts Cash Flow Forecast Errors	23
4.2 Optimism Bias Can Explain Risk Mispricing	24
5 The Required Return of Equity Factors	27
5.1 The Risk Hypothesis is False for Most Factors	28
5.2 Risky Stocks are Small, Volatile, and Distressed	34
6 Can Asset Pricing Models Explain Required Returns?	36
6.1 Candidate Asset Pricing Models	37
6.2 Asset Pricing Models Explain Realized <i>or</i> Required Returns	40
7 Conclusion: A Risk-Return Tradeoff Concealed by Optimism Bias	45
8 Appendix	46
8.1 Proofs	46
8.2 Subjective Risk and Cash Flow Uncertainty	53
8.3 Example of Value Line Investment Report	54

8.4	Subjective Expected Return Calculation	56
8.5	The Subjective Risk Premium is Countercyclical	57
8.6	Subjective Risk and Long-Term Growth Expectations	58
8.7	Equity Factor Information	58
8.8	Machine Learning Predictions via XGBoost	64
8.9	Bootstrap p -Values for Comparing Asset Pricing Models	65
8.10	Robustness	69
2	Is there a Replication Crisis in Finance?	71
1	A Bayesian Model of Factor Replication	78
1.1	Learning About Alpha: The Bayes Case	78
1.2	Hierarchical Bayesian Model	82
1.3	Bayesian Multiple Testing and Empirical Bayes Estimation	87
2	A New Public Data Set of Global Factors	94
3	Empirical Assessment of Factor Replicability	96
3.1	Internal Validity	96
3.2	External Validity	99
3.3	Bayesian Multiple Testing	106
3.4	Economic Significance of Factors	110
4	Conclusion: Finance Research Posterior	115
5	Appendix	116
5.1	Additional Results on Alpha Hacking	116
5.2	Proofs and Lemmas	116
5.3	Empirical Bayes Estimation	123
5.4	Differences in Sample and Factor Construction	125
5.5	Additional Time-Series Results	127
5.6	Economic Benefit of More Powerful Tests	130
5.7	Accounting for Publication Bias	130
5.8	Results by Cluster, Region, and Size	132
5.9	Further Results on the Tangency Portfolio	133
5.10	Cluster Construction	138
5.11	Details on Clusters, Factors, and Countries	139
5.12	Global Factor Data Documentation	153
3	Machine Learning and the Implementable Efficient Frontier	177
1	Model and the Implementable Efficient Frontier	184
1.1	Trading Costs, Net Returns, and Portfolio Growth	185
1.2	Objective Function	186

1.3	The Implementable Efficient Frontier	186
1.4	Empirical Assessments of Portfolios with Trading Costs	187
2	Solution	188
2.1	Optimal Dynamic Portfolio Choice	189
2.2	Implementing the Solution with Machine Learning	192
2.3	Economic Feature Importance	194
3	Benchmarks based on Standard Approaches	195
3.1	Standard Approach: Predicting Returns without T-Costs	195
3.2	Static Transaction Cost Optimization	196
4	Data and Empirical Methodology	197
4.1	Data and Inputs to the Portfolio Choice	197
4.2	Machine Learning Methodology	200
4.3	Portfolio Tuning	201
5	Empirical Results	203
5.1	Out-of-Sample Portfolio Performance	203
5.2	Evidence on the Implementable Efficient Frontier	209
5.3	Economic Feature Importance	212
6	Conclusion	214
7	Appendix	216
7.1	Implementation Details	216
	Computing the Discount Factor m	216
	Implementation Details for Multiperiod-ML	219
7.2	Proofs	219
	Properties of the Implementable Efficient Frontier	219
	Verification Lemma	222
	Properties the Discount Factor m	227
	Proofs of Propositions 12 and 13	230
	Proofs of Propositions 14 and 15	231
	On the Optimality of Portfolio-ML	232
	Proofs related to Economic Feature Importance	234
7.3	Data and Empirical Results	235
	Information of Stock Characteristics (Features)	235
	Portfolio Tuning	235
	Feature Persistence and Importance across Return Horizons	236

Chapter 1

Subjective Risk and Return

Abstract

I use novel data on subjective risk and return expectations to infer investors' required returns. I find that the required compensation for risk is high while the realized compensation for risk is low. This difference arises because cash flow forecasts are systematically too high for risky stocks, which can be explained by investors suffering from optimism bias. The weak link between realized and required returns has two important implications: First, most equity factors have a negative required return despite having a positive realized return. Second, recent empirical asset pricing models explain realized returns well but required returns poorly—while the opposite is true for traditional models like the CAPM.

I am grateful for helpful comments from Jens Dick-Nielsen, Nick Barberis, Paul Ehling (discussant), Marc Eskildsen, Peter Feldhütter, Paul Fontanier, Nicola Gennaioli (discussant), Niels Gormsen, Robin Greenwood, Markus Ibert, Jinoug Jeung (discussant), Bryan Kelly, Stig Lundeby (discussant), Semyon Malamud, Kasper Meisner Nielsen, Lasse H. Pedersen, Alberto Rossi, Alessandro Spina, Julian Terstege, Anders Trolle, Johannes Wohlfart, and seminar and conference participants at Boston University Questrom School of Business, Georgetown University McDonough School of Business, London Business School, London School of Economics, Nova SBE, Rotman School of Management, University College London, Yale School of Management, EFA Doctoral Tutorial 2022, the FRIC Research Retreat 2022, the Bocconi-Tisvildeleje Workshop 2022, the 6th HEC Paris Finance PhD Workshop, and the Young Scholars Nordic Finance Workshop 2022.

Risk-averse investors require a higher return to invest in riskier stocks. Therefore, required returns reflect investors' views on what makes a stock risky and the compensation investors require for taking risks. In this paper, I propose to estimate required returns using observable subjective risk and return expectations.

The standard approach to estimating required returns relies on assumptions about investor beliefs and preferences. For example, required returns follow the CAPM if investors have homogeneous beliefs and mean-variance preferences over portfolio returns (Sharpe, 1964) and the consumption CAPM if a representative investor has time-separable utility over aggregate consumption (Breedon, 1979). However, required returns could also reflect latent state variable risk captured by empirical multifactor models (Fama and French, 1993a, 2015), the ratio of growth options to assets-in-place (Berk et al., 1999), good vs. bad beta (Campbell and Vuolteenaho, 2004), investment-specific technology risk (Kogan and Papanikolaou, 2013, 2014), risk evaluated according to prospect theory (Barberis et al., 2021) and so on. The difficulty with the standard approach is determining which assumptions best capture the behavior of actual investors.

Instead of making relying on assumptions about investor beliefs and preferences, I rely on observable subjective risk and return expectations. Subjective risk ratings reveal which stocks investors perceive as risky but not the risk premium required by investors. I estimate this "subjective risk premium" by regressing subjective expected returns on subjective risk:

$$\text{subjective expected return}^i = \underbrace{\lambda \times \text{subjective risk}^i}_{:= \text{required return}^i} + \epsilon^i, \quad (1.1)$$

where λ is the subjective risk premium and the residual ϵ^i captures perceived mispricing, that is, the subjective view of whether the stock is over- or undervalued. The subjective risk-return relation in (1.1) allow me to infer required returns as the product of subjective risk and the subjective risk premium.

To evaluate whether investors earn the risk compensation they require, I estimate the realized risk premium, λ^{rea} , by regressing realized returns on subjective risk:

$$\text{realized return}^i = \lambda^{rea} \times \text{subjective risk}^i + u^i. \quad (1.2)$$

My first finding is that the subjective risk premium (λ) is high while the realized risk premium (λ^{rea}) is low. This finding implies that investors require a much higher return for investing in risky stocks relative to safe ones, but, nevertheless, risky stocks only deliver slightly higher realized returns. My second finding, is that the gap between required and realized returns arises because cash flow expectations of risky stocks are too optimistic. A

finding which I show theoretically and empirically can be explained by investors suffering from optimism bias.

I make two additional findings based on the required return estimates: First, most equity factors have a negative required return despite having a positive realized return. Second, recent empirical asset pricing models explain realized returns well but required returns poorly. In contrast, traditional models explain required returns well but realized returns poorly.

The Subjective and Realized Risk Premium. Figure 1 shows the relation between subjective expected returns and subjective risk (blue circles). For example, in the top-left corner of Figure 1, I sort stocks into 10 portfolios based on their subjective risk, here measured as the “safety” rank from Value Line, plotted on the x -axis. Value Line is an independent equity research firm, and the data covers 1,700 of the largest US stocks from 1987 to 2021.¹ For each group of stocks, the y -axis shows the average subjective expected return as reported by Value Line. Based on these subjective risk and return expectations, the figure shows the subjective risk-return relation given by the regression line from (1.1). Value Line expects a 20% return from the riskiest stocks but only 8% for the safest, which implies that the subjective risk premium is high.

Figure 1 also displays the relation between subjective risk and average realized returns (red triangles) and the corresponding regression line from (1.2). Focusing again on the top left corner, we see that the realized risk-return relation is positive but weak, as the riskiest stocks only outperform the safest by 3% per year. This result implies that the realized risk premium is low.

These results hold across various subjective risk and return expectations, as seen in the other panels of Figure 1. I use two measures of subjective expected returns—from Value Line and sell-side analysts available in I/B/E/S—and three subjective risk measures—Value Line’s safety rank, Value Line’s market beta estimate, and the subjective risk measure of a log utility investor from Martin and Wagner (2019). Across combinations of risk and return expectations, the subjective risk premium is high and statistically significant, while the realized risk premium is significantly lower. This pattern is even stronger when I use value-weights instead of equal-weights (Figure A5).

Subjective Risk and Cash Flow Optimism. The difference between required and realized returns reflects biased beliefs. To make this point, I use subjective cash flow forecasts from I/B/E/S and Value Line. I find that subjective risk strongly predicts cash flow forecast errors. Specifically, the earnings per share (EPS) forecasts for safe stocks are approximately unbiased, while the EPS forecast for risky stocks is generally much higher

¹Value Line serves both retail and professional investors, with prominent examples of the latter being Warren Buffett, Charlie Munger, and Peter Lynch. They received much academic interest in the 1970s and 1980s for the quality of their recommendations, for example, in Black (1973).

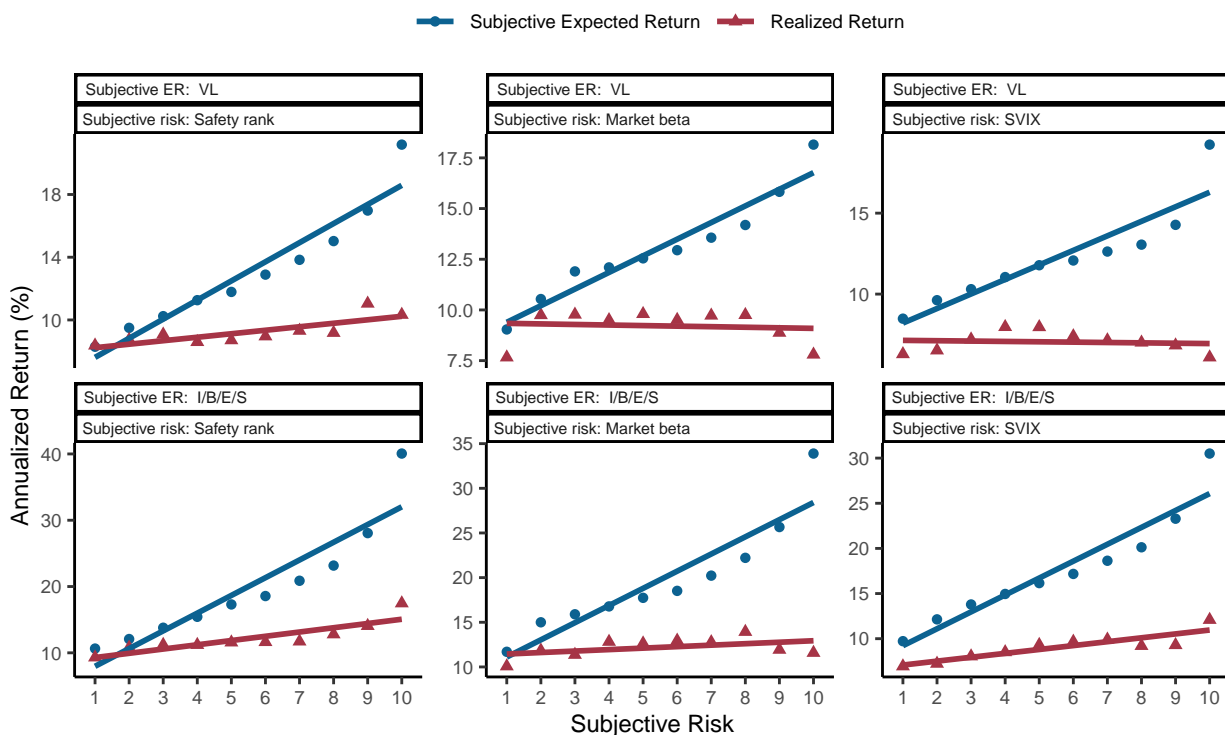


Figure 1. The Subjective Risk Premium is High but the Realized Risk Premium is Low

Note: The figure shows the relation between subjective risk and subjective expected returns and the relation between subjective risk and average realized return for six different combinations of subjective risk and return expectations. Each month, I sort stocks into 10 portfolios according to the subjective risk proxy, where safe (risky) stocks are in portfolio 1 (10). I then compute the average subjective expected return, and the average realized return over the corresponding horizon for each portfolio month. The blue circles refer to the time-series average of the subjective expected return, and the red triangles refer to the time-series average of the realized return. The solid lines show the best linear fit. The name in the first row of the panel title indicates the subjective expected return proxy and the name in the second row shows the subjective risk proxy. The subjective expected return proxy is either a four-year expectation from Value Line or a one-year expectation from I/B/E/S. The subjective risk proxy is either a safety rank from Value Line, a stock’s market beta estimate from Value Line, or the SVIX risk measure from Martin and Wagner (2019).

than the subsequent realization. The high required return of risky stocks is, therefore, muted in realized returns because of irrational cash flow optimism.

The excessive cash flow forecast of riskier stocks suggests that these stocks are over-valued ex-ante. I show theoretically that this “risk mispricing” arises if investors suffer from optimism bias whereby they overweight the probability of good future outcomes (Proposition 4). Optimism bias has a greater effect on risky stocks because they tend to have more uncertain cash flows (thus leaving more room for optimism). Intuitively, excessive optimism is easier to justify for a risky biotech start-up relative to a safe utility stock.

Empirically, I find support for two additional predictions of optimism bias. First, the

average forecast error is positive.² Second, the average forecast error and the forecast error related to subjective risk increase with the forecast horizon. This finding is consistent with my theoretical result that optimism bias has a larger effect when cash flow uncertainty is higher (since cash flow uncertainty typically increases with the cash flow horizon).

Equity Factors. Next, I use the required return estimates to test whether equity factors reflect rational compensation for risk or behavioral mispricing. Under the rational interpretation, factors have a high realized return because they capture risk investors require compensation for bearing. Under the behavioral interpretation, factors capture mispricing. To distinguish between these two alternatives, I test the “risk hypothesis” that a factor’s average realized return is equal to its required return. I derive a null distribution under the risk hypothesis that accounts for the estimation of required returns and that the realized return sample is longer than the one for required returns.

In a sample of 119 factors from Jensen et al. (2022a), I reject the risk hypothesis for 71% to 79% (depending on the risk and returns expectations used to estimate required returns). This high rejection rate primarily reflects that most factors have a weak, or even negative, relation to subjective risk. In particular, only 14%-27% of the factors have a significantly positive required return—a minimal requirement for any risk-based explanation. In contrast, 44%-50% of the factors have a significantly negative required return, meaning that stocks in the long portfolio are subjectively *safer* than stocks in the short portfolio. These results highlight the pitfalls in learning about required returns from realized returns.

For individual factors, I find that the size factor has a high required return while quality and profitability factors have sizeable *negative* required returns. Between these two extremes, the momentum factor and the asset growth factor have a required return close to zero. Focusing on the multivariate drivers of required returns, I find that a typical safe stock is large and profitable, with low return volatility and market beta, while a typical risky stock is small, volatile, and distressed.

Asset Pricing Models. Next, asset pricing models have a dual mandate of predicting realized returns and explaining required returns. They are, however, typically judged solely by how well they perform on the realized return mandate.³

I show that the only case where one model can be optimal for both mandates is when the market is efficient. If the market is inefficient, the optimal model for realized return

²For cash flow expectations from sell-side analysts (I/B/E/S), the positive average forecast error could reflect incentives to produce upwards-biased forecasts to please investment banking clients or to generate trading commission (Kothari, 2001). However, these incentives-related biases are unlikely to explain the positive average forecast error for Value Line since their primary source of income is selling investment research.

³For papers judging asset pricing models solely on their ability to predict realized returns see, for example, Fama and French (1993a), Fama and French (2015), Hou et al. (2015), and Barillas and Shanken (2018).

differs from the optimal model for required returns (Proposition 3).

Empirically, I test the ability of three traditional models, such as the CAPM, and four recent models, such five-factor model from Fama and French (2015), to explain realized and required returns.⁴ I find that the recent empirical models are superior for explaining realized returns. The recent models have R^2 's ranging from 0.26 to 0.45, compared to -0.92 to -0.05 for the traditional models.

For required returns, the ranking is exactly the opposite. The traditional models explain required returns well, with R^2 's ranging from 0.24 to 0.63. The CAPM, in particular, is the best model of required returns for five of six sets of subjective risk and return expectations. In contrast, the recent empirical models tend to imply a high expected return for portfolios with a low required return, and, as a result, their R^2 's are deeply negative. The recent models have improved our ability to explain realized returns but not required returns.

Related literature. To my knowledge, I am the first to study subjective risk and return expectations in the cross-section of stocks. For the overall stock market, Nagel and Xu (2022b) finds a positive relationship between subjective risk and return expectations, that is, a positive subjective risk-return tradeoff. In contrast, when surveying households, Jo et al. (2022) finds a negative subjective risk-return tradeoff across asset classes. I contribute to this literature by showing a strongly positive risk-return tradeoff in the cross-section of stocks.

At the stock level, Lui et al. (2007) show that risk ratings from Salomon Smith Barney (SSB, now Citigroup) primarily depend on idiosyncratic volatility, size, book-to-market, and leverage.⁵ At the market level, Lochstoer and Muir (2022) find that subjective volatility expectations underreact to news initially, followed by a delayed overreaction. I contribute to this literature by using subjective risk to infer required returns.

Brav et al. (2005) show that return expectation from Value Line and I/B/E/S increase in market beta, decrease in firm size, and is unrelated to the book-to-market ratio. Similarly, Engelberg et al. (2020) show that return expectations and stock recommendations from I/B/E/S are negatively related to the average equity factors. I contribute to this literature by showing that the required return of most factors is negative.

Gormsen and Huber (2022) show that the cost of capital used by CFOs closely follow the CAPM, while Berk and Van Binsbergen (2016) show that mutual fund flows are con-

⁴The traditional models are the CAPM, the Consumption CAPM, the Fama-French three-factor model, and the Carhart four-factor model, while the recent models are the five-factor model from Fama and French (2015), the q -model from Hou et al. (2015), the model with mispricing factors from Stambaugh and Yuan (2017), and the model with behavioral factors from Daniel et al. (2020). I test the models on the 3×119 high-middle-low characteristics sorted portfolios that underlies the 119 equity factors.

⁵Lui et al. (2007) also considers the safety rank from Value Line and finds that it has a correlation of 0.61 with the risk rating from SSB and a correlation of 0.65 with a risk rating from Merrill Lynch. These high correlations provide external validity for the safety rank as a measure of subjective risk.

sistent with investors using the CAPM. I find that the CAPM is superior for explaining direct proxies of investors’ required returns. Further, I show that irrational cash flow optimism can explain why the CAPM fails to explain realized returns despite its widespread use.

I also contribute to the literature that uses subjective expectation in asset pricing. For individual stocks, most papers use data on subjective cash flow expectations (e.g., La Porta (1996), Engelberg et al. (2018), Bordalo et al. (2019), and Bouchaud et al. (2019)), while the use of subjective return expectations is more common at the market level (e.g., Vissing-Jorgensen (2003), Amromin and Sharpe (2014), Greenwood and Shleifer (2014), and Dahlquist and Ibert (2021)). I contribute to this literature by jointly studying subjective risk, return, and cash flow expectations for stocks, factors, and asset pricing models.

1 Theory

1.1 Setup

I consider an economy with stocks indexed by $i = 1, \dots, N$ and discrete time periods indexed by $t = 1, 2, \dots$. The one-period total return of a stock is $r_{t+1}^i + r_{t+1}^f$, where r_{t+1}^i is the stock’s excess return and r_{t+1}^f is the risk-free rate. I denote objective expectations by E and subjective expectations by \tilde{E} . The subjective expectations are from the perspective of a single investor.⁶

The investor (implicitly or explicitly) ranks stocks according to their subjective risk, s_t^i . The investor is risk-averse and therefore requires an additional return of $\lambda_t > 0$ (the “subjective risk premium”) for each additional unit of risk. Hence, the required return for investing in a stock from the perspective of the investor is:

$$\text{Required Return}_t^i = \lambda_t s_t^i. \tag{1.3}$$

The investor also computes an expected return for each stock, $\tilde{E}_t[r_{t+1}^i]$. The difference between the required and expected return captures subjective mispricing, $\tilde{b}_t^i = \lambda_t s_t^i - \tilde{E}_t[r_{t+1}^i]$. The investor views a stock as undervalued if the expected return is higher than the required return and overvalued if the expected return is lower than the required return. Re-arranging the expression for the subjective mispricing, the subjective expected return is:

$$\tilde{E}_t[r_{t+1}^i] = \lambda_t s_t^i - \tilde{b}_t^i. \tag{1.4}$$

⁶With J investors, there is, potentially, J different subjective expectations. A key question is whether the expectation of a specific investor accurately reflects the remaining $J - 1$. Empirically, I tackle this issue by using six different proxies for subjective expectations.

The investor determines the required return, so there is no distinction between the objective and subjective required return. However, the subjective mispricing reflects the investor’s realized return forecast, and this forecast could be biased. Let $E_t[r_{t+1}^i]$ be the objectively correct forecast of the realized return—the objective expected return. The objective mispricing of a stock is then, $b_t^i = \lambda_t s_t^i - E_t[r_{t+1}^i]$.⁷

The relation between subjective risk and objective mispricing is key to understanding the difference between realized and required returns. Therefore, without loss of generality, I regress objective mispricing on subjective risk across stocks,

$$b_t^i = \gamma_t^0 + \gamma_t^1 s_t^i + u_t^i, \quad (1.5)$$

where γ_t^0 and γ_t^1 is the coefficient from a cross-sectional regression of objective mispricing on subjective risk, and u_t^i is a residual. The relation between risk and mispricing is captured by γ_t^1 , while u_t^i reveals mispricing unrelated to risk (“non-risk mispricing”).

1.2 Results

The Subjective and Required Risk Premium

In the empirical analysis, I estimate the subjective risk premium, λ_t , by regressing subjective expected returns on subjective risk,

$$\tilde{E}_t[r_{t+1}^i] = a_t + \hat{\lambda}_t s_t^i + \epsilon_t^i, \quad (1.6)$$

where a_t is the intercept of the regression, $\hat{\lambda}_t$ is my estimate of λ_t , and ϵ_t^i is a residual. In addition, I estimate the realized risk premium by regressing objective expected returns on subjective risk,⁸

$$E_t[r_{t+1}^i] = a_t^{rea} + \hat{\lambda}_t^{rea} s_t^i + \epsilon_t^{rl,i}, \quad (1.7)$$

where $\hat{\lambda}_t^{rea}$ is the realized risk premium estimate.

Proposition 1 shows when $\hat{\lambda}_t$ is a good proxy for the subjective risk premium, λ_t , and how this subjective risk premium differs from the realized risk premium, $\hat{\lambda}_t^{rea}$ (all proofs are in appendix 8.1).

⁷To understand b and \tilde{b} , consider a stock with a price today of 10 and a liquidating cash flow in one year. The investor expects the cash flow to be 12, the risk-free rate is 0%, and the investor requires $\lambda_t s^i = 10\%$ for investing in the stock. The investor’s subjective expected return is $\tilde{E}_t[r_{t+1}^i] = 12/10 - 1 = 20\%$ and the subjective mispricing is $\tilde{b}_t^i = -10\%$. That is, the investor views the stock as 10% undervalued. Suppose the investor is biased and the objective cash flow expectation is 11. The objective expected return is then $E_t[r_{t+1}^i] = 11/10 - 1 = 10\%$ and the objective mispricing is $b_t^i = 10\% - 10\% = 0\%$.

⁸In practice, I regress realized returns on subjective risk, but the coefficients reveal the relationship between subjective risk and objective expected returns.

Proposition 1 (The Subjective and Realized Risk Premium) *The subjective risk premium estimate, $\hat{\lambda}_t$ from (1.6), depends on the true subjective risk premium and the relation between subjective risk and subjective mispricing,*

$$\hat{\lambda}_t = \lambda_t - \frac{\overline{\text{Cov}}(s_t^i, \tilde{b}_t^i)}{\overline{\text{Var}}(s_t^i)}, \quad (1.8)$$

where $\overline{\text{Var}}$ and $\overline{\text{Cov}}$ is the cross-sectional variance and covariance, respectively. The realized risk premium estimate, $\hat{\lambda}_t^{\text{rea}}$ from (1.7), depends on the true subjective risk premium and the relation between subjective risk and objective mispricing,

$$\hat{\lambda}_t^{\text{rea}} = \lambda_t - \frac{\overline{\text{Cov}}(s_t^i, b_t^i)}{\overline{\text{Var}}(s_t^i)} = \lambda_t - \gamma_t^1, \quad (1.9)$$

where the last equality uses the definition of γ_t^1 from (1.5).

Equation (1.8) in Proposition 1 shows that $\hat{\lambda}_t$ is an unbiased estimate of λ_t if subjective risk and subjective mispricing are uncorrelated. If, for example, the investor on average views riskier stocks as overvalued, $\overline{\text{Cov}}(\tilde{b}, s_t^i) > 0$, then the estimated risk compensation is lower than the true compensation. Therefore, I can recover the investor's required returns from subjective risk and return expectations if there is no systematic (linear) relation between subjective risk and subjective mispricing. In addition, equation (1.9) in Proposition 1 shows that the difference between the required and realized compensation for risk depends on γ_t^1 . Empirically, I find evidence of $\gamma_t^1 > 0$, implying that the realized compensation for risk is lower than investors require.

Equity Factors

Next, I consider an equity factor that goes long/short stocks according to some underlying stock characteristics. The return of this portfolio is $r_{t+1}^L - r_{t+1}^S = r'_{t+1} \pi_t^L - r'_{t+1} \pi_t^S$, where π^S and π^L is the portfolio weights in the long and short portfolio, respectively. For simplicity, I consider a dollar neutral long-short factor where $\sum_{i \in L} \pi_{i,t}^L = \sum_{i \in S} \pi_{i,t}^S = 1$, and I sign the factor such that the long portfolio has a higher objective expected return. This setup is inspired by the extensive literature on equity factors. Rational theories explaining why a factor works assume that the factor's realized return is equal to its required return. The following proposition shows when this assumption fails:

Proposition 2 (Equity Factors) *The objective expected return of a dollar neutral equity factor depends on its required return, the relation between risk and objective mispricing, and the factors exposure to non-risk mispricing,*

$$\mathbb{E}_t[r_{t+1}^L - r_{t+1}^S] = (\lambda_t - \gamma_t^1)(s_t^L - s_t^S) - (u_t^L - u_t^S), \quad (1.10)$$

where $s_t^k = \sum_{i \in k} \pi_t^i s_t^i$ is the weighted subjective risk of stocks in portfolio j , and $u_t^k = \sum_{i \in k} \pi_t^i u_t^i$ is the weighted non-risk mispricing.

Proposition 2 shows that the objective expected return of an equity factor is different from its required return ($\lambda_t(s_t^L - s_t^S)$) if the required and realized risk compensation differ, $\gamma_t^1 \neq 0$, or if the factor captures exposure to non-risk mispricing. Since $\gamma_1 > 0$ empirically, the proposition imply that the realized return of true risk factors is lower than their required returns all else equal. As such, we could have missed factors that truly matter for required returns or have downplayed the importance of certain risk factors. Similarly, a factor's exposure to non-risk mispricing can disguise its required return. For example, factors can have a high realized return even if stocks in the long and short portfolio have similar risks if the exposure to non-risk mispricing is sufficiently high.

Asset Pricing Models

Asset pricing models are typically judged by their ability to predict realized returns, or, said differently, by how well they align with objective expected returns (e.g., Fama and French (1993a, 2015), Hou et al. (2015), and Barillas and Shanken (2018)). At the same time, fundamental investors and corporate managers use required returns from asset pricing models to value assets and for other corporate decisions. I want to understand whether empirical models designed to explain realized returns should be used for required returns.

I consider two types of empirical models. The first model uses a pricing factor, f_t^1 , where a stock's loading with respect to this factor perfectly maps to subjective risk, $\overline{\text{Cov}}_t(r_{t+1}^i, f_{t+1}^1) / \overline{\text{Var}}_t(f_{t+1}^1) = s_t^i$, but estimates the factor premium empirically (freely) by regressing objective expected returns on f_t^1 ,

$$E_t^{\text{free}}[r_{t+1}^i] = \kappa_t^0 + \kappa_t^1 s_t^i, \quad (1.11)$$

where κ_0 and κ_1 are the coefficients from regressing objective expected returns on subjective risk. This model is inspired by empirical implementations of theoretical models that often estimate risk premiums freely instead of using the values implied by the theory. The second empirical model adds an additional pricing factor, f_t^2 , where a stock's loadings with respect to this factor, c_t^i , is correlated with non-risk mispricing, $\overline{\text{Cov}}(c_t^i, u_t^i) \neq 0$. That is, stocks with a high (absolute) loading on f_t^2 tend to have a high (absolute) non-risk mispricing. The model-implied expected return is,

$$E_t^{\text{multifactor}}[r_{t+1}^i] = \kappa_t^0 + \kappa_t^1 s_t^i + \kappa_t^2 c_t^i, \quad (1.12)$$

where κ_2 is the coefficient from regressing objective expected returns on a constant and

loadings from the two factors. For simplicity, I assume that the loadings on the new factor are mean zero and uncorrelated with subjective risk. This model is meant to resemble empirical multifactor models such as the five-factor model of Fama and French (2015) and the q -model of Hou et al. (2015). I compare the empirical models to two benchmark models. The first benchmark model perfectly matches required returns,

$$\mathbb{E}_t^{\text{req}}[r_{t+1}^i] = \lambda_t s_t^i, \quad (1.13)$$

while the second model perfectly matches objective expected returns,

$$\mathbb{E}_t^{\text{obj}}[r_{t+1}^i] = \mathbb{E}_t[r_{t+1}^i]. \quad (1.14)$$

To evaluate a model's ability to explain realized returns, I define the objective pricing error for any model m as,

$$\alpha_m^2 = \mathbb{E} \left[\left(\mathbb{E}_t[r_{t+1}^i] - \mathbb{E}_t^m[r_{t+1}^i] \right)^2 \right], \quad (1.15)$$

where $\mathbb{E}_t^m[r_{t+1}^i]$ is the model-implied expected return. The lower the objective pricing error, the better a model fulfills the realized return mandate. This metric is closely related to the ‘‘GRS’’ test of Gibbons et al. (1989), where the null hypothesis is that α_m^2 is zero. Similarly, to investigate a model's ability to explain required returns, I define the subjective pricing error as:

$$\tilde{\alpha}_m^2 = \mathbb{E} \left[\left(\lambda_t s_t^i - \mathbb{E}_t^m[r_{t+1}^i] \right)^2 \right]. \quad (1.16)$$

The lower the subjective pricing error, the better a model fulfills the required return mandate.

The empirical asset pricing models aim to fulfill the return mandate by minimizing the objective pricing error from (1.15). As such, the models try to approximate the objective model from (1.14). In contrast, a model that fulfills the required return mandate should approximate the model from (1.13). The following proposition shows the extent to which empirical asset pricing models can fulfill both mandates simultaneously:

Proposition 3 (Asset pricing models) *The relative ranking of the four models for explaining realized returns is*

$$\alpha_{\text{req}}^2 \geq \alpha_{\text{free}}^2 \geq \alpha_{\text{multifactor}}^2 \geq \alpha_{\text{obj}}^2 = 0,$$

while the relative ranking for required returns is in the opposite order

$$\tilde{\alpha}_{obj}^2 \geq \tilde{\alpha}_{multifactor}^2 \geq \tilde{\alpha}_{free}^2 \geq \tilde{\alpha}_{req}^2 = 0,$$

Proposition 3 shows that, for the four models, their relative ability to explain realized returns is exactly opposite to their relative ability to explain required returns. As such, when empirical models get better at explaining realized returns, they get worse at explaining required returns. Conversely, a model that tries to explain required returns has an inferior realized pricing ability. This result implies that one model cannot simultaneously be optimal for explaining realized and required returns. The only exception to this claim is the special case where objective expected returns *only* reflect required returns, $E_t[r_{t+1}^i] = \lambda_t s_t^i$. In that case, all four models coincide and have a zero pricing error according to both pricing metrics.

Optimism Bias

Finally, I show that an investor with an “optimism bias” makes larger mistakes when forecasting the cash flows of riskier stocks. Optimism bias refers to the tendency of most people to overweight the probability of good future outcomes. For example, Sharot (2011) estimates that around 80% of forecasters have an optimism bias while the remaining 20% have a pessimism bias.

Consider an investor j that wants to forecast the cash flow of stock i denoted x_i . The dividend has an expectation of θ_i and a variance of ω_i . The prior distribution of θ_i is normal with a common mean of μ_0 and a stock-specific variance which is proportional to its cash flow variance: $\tau_0^2 \omega_i^2$, where $\tau_0 > 0$ is a constant.

Investors observe two public signals for each stock which are independently drawn from a normal distribution with a mean of θ_i and variance $\tau_1^2 \omega_i^2$ where $\tau_1 > 0$ is a constant. Without loss of generality, I define v_i^{max} as the highest of the two signals and v_i^{min} as the lowest.

Investors use the two signals to infer the value of θ_i . The rational (Bayesian) approach is to shrink the prior mean towards the equal-weighted average of the two signals:

$$E[\theta_i | v_i^{max}, v_i^{min}] = \mu_0 + \delta \left(\frac{1}{2} v_i^{max} + \frac{1}{2} v_i^{min} - \mu_0 \right), \quad (1.17)$$

where $\delta = \frac{\omega_i^2 \tau_0^2}{\omega_i^2 \tau_0^2 + \omega_i^2 \tau_1^2 / 2} = \frac{\tau_0^2}{\tau_0^2 + \tau_1^2 / 2}$ is the shrinkage constant, which is the same for all stocks.

However, I allow for the possibility that actual investor inference is non-Bayesian. In particular, I assume that the investor puts κ^j weight on v_i^{max} and $1 - \kappa_j$ weight on v_i^{min} :

$$\tilde{E}^j[\theta_i | v_i^{max}, v_i^{min}] = \mu_0 + \delta (\kappa^j v_i^{max} + [1 - \kappa_j] v_i^{min} - \mu_0). \quad (1.18)$$

An investor with an optimism bias puts too much weight on the good signal (v_i^{max}), which is captured by $\kappa^j > 0.5$.

To characterize the error made by the investor, I define the bias in an investor’s subjective cash flow expectation as $b_i^j = \tilde{E}^j[\theta_i|v_i^{max}, v_i^{min}] - E[\theta_i|v_i^{max}, v_i^{min}]$. This bias depends on the two random signals, but the next proposition shows its expected value:

Proposition 4 (The effect of optimism bias increases in cash flow uncertainty)

The expected bias is,

$$\tilde{E}[b_i^j] = c(\kappa_j - 0.5)\omega_i, \tag{1.19}$$

where $c \geq 0$ is constant across stocks. For an investor with an optimism bias, $\kappa^j > 0.5$, the expected bias increases in the stock’s cash flow volatility, ω_i .

Proposition 4 shows that investors who suffer from an optimism bias make larger mistakes for stocks with higher cash flow uncertainty. The reason is that stocks with more uncertain cash flows leave more room for optimism. To see this, consider first an optimistic investor trying to forecast the cash flows of a firm with low cash flow uncertainty, such as a utility firm. Utility firms tend to have regulated pricing structures and long-term contracts, so the cash flow in “good” (v_i^{max}) and “bad” (v_i^{min}) states are similar. The forecast of the optimistic investors is, therefore, similar to that of the rational investor simply because the low cash flow uncertainty leaves limited room for optimism. In fact, in the extreme case where the cash flow is known (say, for a US government bond), the forecast of the optimistic investor is equal to that of the rational investor.

In contrast, consider an optimistic investor trying to forecast the cash flows of a firm with high cash flow uncertainty, such as Tesla. In some states of the world, electric vehicles completely replace traditional cars, with Tesla as the market leader. In other states, Tesla goes bankrupt. As a result, Tesla leaves plenty of room for optimism meaning that the expected bias of an optimistic investor is high.

Appendix 8.2 shows that cash flow uncertainty is strongly increasing it stock risk. As a result, Proposition 4 predicts that cash flow mistakes should be larger for riskier stocks if the forecasters suffer from an optimism bias. Hence, optimism bias predicts that the gap between the subjective expected return and the average realized return of a stock is increasing in its subjective risk, consistent with the findings in Figure 1.

2 Data

Subjective Risk and Return Expectations

I use subjective risk and return expectation from various sources, to ensure that my results are not specific to a particular set of expectation. In particular, I use three proxies for

subjective risk and two proxies for subjective expected returns from three different data sources summarized in Table I.

Table I. Subjective Risk and Return Proxies

Name	Notation	Provider	Period	#Stocks
Subjective risk				
Safety rank	$s_t^{VL,i}$	Value Line	1990/07 - 2021/12	1,461
Market beta	β_t^i	Value Line	1990/07 - 2021/12	1,461
SVIX	$SVIX_t^i$	Martin and Wagner (2019)	1996/01 - 2013/09	409
Subjective expected returns				
VL ER	$\tilde{E}_t^{VL}[r_{t,t+48}^i]$	Value Line	1990/07 - 2021/12	1,439
IBES ER	$\tilde{E}_t^{IBES}[r_{t,t+12}^i]$	I/B/E/S	1999/06 - 2021/12	3,789

Note: This table shows the subjective risk and subjective expected return proxies I use throughout the paper. The column “Name” shows proxies name, “Notation” the notation I use to refer to the proxy, “Provider” the data provider, “Period” the period where the proxy is available, and “#Stocks” the median number of stocks the proxies cover across months.

The most comprehensive and novel data comes from Value Line, an independent equity research firm founded in 1931 that currently employs 70+ equity analysts. Value Line’s flagship product is the weekly publication of the “Value Line Investment Survey,” which contains summary statistics for 1,700 of the largest US stocks and an in-depth analysis of 130-140 of these stocks. Each stock gets an in-depth review once per quarter or if something material happens to the underlying company. Appendix 8.3 presents an example of a report on Apple. I have access to the Value Line data starting in 1987, but the subjective risk and return expectations are available from 1990.

Value Line’s customers range from individual investors who pay an annual subscription fee of \$795 for basic services to professional investors who pay more than \$100,000 annually.⁹ Prominent investors who have used Value Line include Warren Buffett and Charlie Munger (CNBC, 1998) and Peter Lynch (Lynch and Rothchild, 2000). Value Line also received considerable academic attention in the 1970s and 1980s for the quality of their stock recommendations (the “timeliness rank”). For example, Black (1973) shows that Value Line had statistically significant stock-picking skills and argued that this provided hope for active managers.

The first proxy for subjective risk is the primary risk measure from Value Line called the “safety rank.” The safety rank ranges from 1 for the safest stocks to 5 for the riskiest. The safety rank is derived by taking an average of a stock’s rank with respect to two sub-

⁹Value Line’s annual report from 2021, p. 15: “Value Line serves primarily individual and professional investors in stocks, who pay mostly on annual subscription plans, for basic services or as much as \$100,000 or more annually for comprehensive premium quality research, not obtainable elsewhere.”

ratings, namely a “price stability index” and a “financial strength rating.”¹⁰ To avoid losing information from putting stocks into discrete bins, I re-define the safety rank as the average cross-sectional rank of price stability and financial strength. I transform this measure into a cross-sectional percentile rank such that 0 is the safest stock in a month and 1 is the riskiest. The average cross-sectional correlation between the safety rank and the binned version of subjective risk is 92%. I denote this re-scaled safety rank as $s_t^{VL,i}$.

One concern is that the safety rank captures the total risk of a stock, whereas some investors set required returns based on the systematic risk of a stock. Fortunately, a Value Line report also includes a measure of systematic risk in the form of a stock’s market beta, β_t^i . In principle, investors could care about systematic risk besides market risk (e.g., inflation risk, consumption risk, etc.). However, the fact that Value Line does not include exposure to other systematic factors suggests that their customers view market beta as a sufficient measure of systematic risk.

The third and final subjective risk proxy is a risk measure from Martin and Wagner (2019).¹¹ The measure is based on the risk-neutral variance of a stock, as computed from option prices. Martin and Wagner (2019) shows that this measure captures subjective risk as perceived by an unconstrained log utility investor. The data is available for S&P500 stocks from 1996 to 2013, and I denote the measure as $SVIX_t^i$. Consistent with the construction of $s_t^{VL,i}$, I transform β_t^i and $SVIX_t^i$ into cross-sectional percentile ranks.

The first proxy of subjective expected returns is from Value Line. The main input is a high and low price target issued over a three-to-five-year horizon, which I assume is realized after four years. I take the simple average of the high and low targets to arrive at the expected price. I combine this price target with dividend expectations from Value Line and the current price from CRSP to compute a stock’s expected total return.¹² The full procedure is described in Section 8.4. To obtain the four-year expected excess return, I need to subtract the four-year risk-free rate. As a proxy, I compound the annualized market-implied yield on a constant five-year maturity US treasury bond provided by the St. Louis FED at <https://fred.stlouisfed.org/series/DGS5>. I denote the expected return from Value Line by $\tilde{E}_t^{VL}[r_{t,t+48}^i]$.

The second proxy of subjective expected return is from sell-side analysts, which I

¹⁰In their database manual, Value Line explains the safety rank as: “Safety is a measurement of the total risk of a stock. Total risk is different from Beta. The latter measures only the extent to which a stock normally responds to changes in the trend and level of the market as a whole. The Safety Rank is a more comprehensive measure of risk including all those factors peculiar to the company’s business such as its financial condition, management competence, etc. The Safety Rank is derived by averaging two variables: 1) the stock’s index of Price Stability, and 2) the Financial Strength rating of the company” (ValueLine, 2021).

¹¹I thank Ian Martin and Christian Wagner for making their data available. I use the measure denoted by $SVIX_{i,t}^2$ in Martin and Wagner (2019).

¹²I require that the price from CRSP is within 5% of the price reported from Value Line. This procedure eliminates 1.1% of the relevant observations.

access through I/B/E/S. Specifically, I use the median one-year price target from the I/B/E/S consensus file, which I supplement with one-year dividend expectations from I/B/E/S and the current price from CRSP. I describe the full procedure in appendix 8.4. I subtract the one-year market-implied yield on a constant one-year maturity US treasury bond available at <https://fred.stlouisfed.org/series/DGS1>. The return expectations from I/B/E/S are available for a large cross-section of stocks from 1999 to 2021, and I denote the resulting measure by $\tilde{E}_t^{IBES}[r_{t,t+12}^i]$.

Subjective Cash Flow Expectations

I acquire subjective cash flow expectations from Value Line and I/B/E/S, primarily reflecting earnings per share (EPS) forecasts. From I/B/E/S, I extract the long-term growth in EPS forecasts `EPS (fpi` of 0), annual EPS forecasts over the next two fiscal years (`fpi` of 1 and 2), and quarterly EPS forecasts over the next four fiscal quarters. I always use the median forecast from the unadjusted consensus file.

From Value Line, I obtain annual EPS forecasts over the next two fiscal years and an EPS forecast over a three-to-five year horizon, which I assume reflects a forecast horizon of four years. In addition, I estimate Value Line’s EPS forecast in fiscal year three by linear interpolation between the two and four-year forecasts and the EPS forecast in fiscal year five by linear extrapolation from the same two points.¹³

For each firm-fiscal year pair, I only retain the first EPS forecast issued at least 45 days and no more than 180 days after the announcement of the previous fiscal year. This gap ensures that the forecast has had time to reflect the previous fiscal year’s information. I get earnings announcement dates from I/B/E/S.

Finally, I compute forecast errors using the EPS realization from the I/B/E/S unadjusted “actuals” file, and I winsorize all forecast errors at the top/bottom 1% to limit the influence of outliers.

Stock Returns and Characteristics

I obtain price and return data from CRSP and accounting data from Compustat. I restrict the sample to ordinary common stocks (`shrcd` of 10, 11, and 12 in CRSP) listed on NYSE, AMEX or Nasdaq (`exchcd` 1, 2, and 3 in CRSP). To create stock characteristics, I use the code from Jensen et al. (2022a) available at <https://github.com/bkelly-lab/ReplicationCrisis>. I compute multi-period returns by compounding monthly returns (with the 1-month risk-free rate added) and subtracting the risk-free rate over the multi-period horizon. If a stock delists, I incorporate its delisting return and assume that the

¹³Specifically, I estimate Value Line’s three- and five-year EPS forecast as $\tilde{E}[eps_{t+h}^i] = \tilde{E}[eps_{t+2}^i] + (\tilde{E}[eps_{t+4}^i] - \tilde{E}[eps_{t+2}^i])(h - 2)$ where $\tilde{E}_t[eps_{t+h}^i]$ is Value Line’s EPS forecast for fiscal year $t + h$.

stock earns the median returns across all stocks in CRSP in the remaining horizon. This procedure avoids any look-ahead bias.

Asset Pricing Models

For asset pricing models with tradable pricing factors, I download these factors from the respective authors' websites: I obtain data for the CAPM, the three-factor model from Fama and French (1993a), the five-factor model from Fama and French (2015), and the four-factor model from Carhart (1997) from Kenneth French's data library at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. The Hou et al. (2015) factors are from <https://global-q.org>, the Daniel et al. (2020) factors are from <https://sites.google.com/view/linsunhome>, and the Stambaugh and Yuan (2017) factors are from <https://finance.wharton.upenn.edu/~stambaug/>. The Stambaugh and Yuan (2017) factors are only available until 2016, so I extend these data series by creating the SMB, PERF, and MGMT factors manually using the `mispricing_perf` and `mispricing_mgmt` characteristics from Jensen et al. (2022a), and proxy for MKT using the Fama-French market factor. For the consumption CAPM, I measure consumption per capita using PCE (consumption) divided by POP (US population) extracted from the data library of the St. Louis FED, <https://fred.stlouisfed.org/>.

Combining Data Sets

The identifier in the Value Line data is a stock's eight-digit CUSIP and exchange ticker. I merge this data with CRSP using the `crsp.stocknames` table on WRDS. I match securities on historical CUSIPs; if this match fails, I use header CUSIPs and historical tickers. In addition, I require that the stock price from Value Line is within 5% of the most recent stock price in CRSP. In total, I match 95% of the Value Line observations to CRSP. Next, I use the `wrdsapps.ibcrsphist` table from WRDS to link CRSP and I/B/E/S data and the `crsp.ccmxpf_lnkhist` table from WRDS to link CRSP and Compustat.

3 The Subjective and Realized Risk Premium

In this section, I first study the relation between subjective risk and subjective expected returns to infer the subjective risk premium that, together with subjective risk, determines required returns. Next, I study the relation between subjective risk and realized returns to infer the realized risk premium. Finally, I study the relation between subjective risk and subjective cash flow errors to understand the impact of mispricing on the difference between the subjective and realized risk premium. This section is closely related to Proposition 1.

3.1 The Subjective Risk Premium is High

The subjective risk premium is the compensation investors require for taking an additional unit of risk, as seen in (1.3). To recover the subjective risk premium, I regress subjective expected returns on subjective risk,¹⁴

$$\tilde{E}_t[r_{t,t+h}^i] = a + \hat{\lambda}s_t^i + \epsilon_t^i, \quad (1.20)$$

where $\tilde{E}_t[r_{t,t+h}^i]$ is one of the two subjective expected return proxies, s_t^i is one of the three subjective risk proxies, a is an intercept, ϵ_t^i is a residual, and $\hat{\lambda}$ the subjective risk premium estimate. Proposition 1 shows that this estimate is unbiased if subjective risk and subjective mispricing is uncorrelated.

Table II presents the estimates for each of the six combinations of a subjective risk and subjective expected return proxy. The impact of subjective risk on subjective expected return is strongly positive and highly significant across all proxies. For example, with $\tilde{E}_t^{VL}[r_{t,t+48}^i]$ and $s_t^{VL,i}$ as proxies (column 1), moving from the safest stock ($s_t^{\text{safe}} = 0$) to the riskiest ($s_t^{\text{risky}} = 1$), increases the four-year subjective expected return from 28% to 103%. Said differently, the required annual return of the safest stock is approximately 6% versus 19% for the riskiest stock.

Interestingly, $s_t^{VL,i}$ has the strongest association with subjective expected returns among the three subjective risk proxies. This finding reduces the concern that $s_t^{VL,i}$ captures total risk rather than systematic risk and is therefore unimportant for investors. On the contrary, $s_t^{VL,i}$ is highly correlated with investors' subjective return expectations consistent with the view that it matters for required returns.

In the remainder of the paper, I use a stock's subjective risk and the coefficients from Table II to estimate a stock's required returns. Said differently, a stock's required return is the constant, a , plus the subjective risk premium, $\hat{\lambda}$, times the stock's subjective risk, s_t^i . For each stock, I obtain six different required return estimates depending on the combination of the subjective risk and subjective expected return proxies as seen from the six columns in Table II.

Subjective Expected Returns Mostly Reflect Subjective Mispricing

Subjective expected returns can reflect required returns and subjective mispricing as seen in (1.4). Further, if subjective risk and subjective mispricing are uncorrelated, the predicted part of (1.20) captures the impact of required returns while the residuals capture subjective mispricing.

¹⁴The specification in (1.20) differs from the theoretical specification in (1.3) to account for the arbitrary scale of the subjective risk proxies. In particular, the subjective risk proxies range between 0 and 1, but a rating of 0 does not imply that the stock is risk-free.

Table II. The Subjective Risk Premium is High

	$\tilde{E}_t^{VL}[r_{t,t+48}^i]$			$\tilde{E}_t^{IBES}[r_{t,t+12}^i]$		
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.28 (23.24)	0.38 (26.96)	0.34 (16.91)	0.07 (13.27)	0.10 (17.34)	0.08 (13.46)
$s_t^{VL,i}$	0.75 (25.82)			0.26 (14.90)		
β_t^i		0.53 (17.75)			0.19 (12.28)	
SVIX $_t^i$			0.53 (9.74)			0.19 (9.12)
Observations	504,101	504,095	81,397	375,267	375,267	70,370
Adj. R 2	0.08	0.04	0.06	0.05	0.03	0.05
Adj. R 2 (time-varying)	0.24	0.21	0.33	0.27	0.23	0.38

Note: The table estimates the subjective risk premium, λ , by regressing subjective expected return on subjective risk. The subjective expected return proxies are a four-year forecast from Value Line ($\tilde{E}_t^{VL}[r_{t,t+48}^i]$), and a one-year forecast from I/B/E/S ($\tilde{E}_t^{IBES}[r_{t,t+12}^i]$). The subjective risk proxies are the safety rank from Value Line ($s_t^{VL,i}$), the market beta estimate from Value Line, and the SVIX $_t^i$ measure from Martin and Wagner (2019). The subjective risk proxies are scaled to lie between 0 and 1 each month, so the coefficients show the change in the dependent variable by moving from the safest (0) to the riskiest (1) firm. The number in parenthesis refers to the t -statistic of the coefficient based on standard errors clustered by firm and quarter. The ‘‘Observations’’ row shows the number of monthly stock observations used to estimate the coefficients and the ‘‘Adj. R 2 ’’ row shows the adjusted R 2 of the regression. The ‘‘Adj. R 2 (time-varying)’’ row shows the adjusted R 2 with the parameters estimated separately in each month, that is, with time-varying parameters.

As a result, the R 2 from Table II shows the fraction of subjective expected return variance explained by required returns. Using the specification in (1.20), the variance explained by required returns is between 3% and 8%, which implies that subjective mispricing explains the remaining 92% to 97%. This finding suggests that views on mispricing are the key driver of variation in subjective expected returns. It also highlights the danger of treating subjective return expectations as required returns.

The specification in (1.20) assumes the mapping from subjective risk to required returns are fixed over time. However, it seems reasonable to assume that the subjective risk premium is time-varying. Therefore, the row labeled ‘‘Adj. R 2 (time-varying)’’ shows the adjusted R 2 from the regression,

$$\tilde{E}_t[r_{t,t+h}^i] = a_t + \hat{\lambda}_t s_t^i + \epsilon_t^i, \quad (1.21)$$

where the coefficients are allowed to vary over time.

The R 2 from the regression in (1.21) is much higher than with fixed parameters ranging

from 21% to 38%. Subjective mispricing is still the most important driver of subjective expected returns, but the importance of required return rises when accounting for the time variation in the subjective risk premium.

In the remainder of the paper, I estimate required returns using the specification in (1.20), that is, with fixed parameters over time. Proposition 1 shows that subjective risk and subjective mispricing must be uncorrelated to recover an unbiased estimate of the subjective risk premium. The choice of using fixed parameters reflects my judgment that this zero correlation condition is unlikely to hold exactly in each period. For example, Value Line could believe that risky stocks are overvalued in some periods and undervalued in others, which would cause the estimated subjective risk premium to be too high and low, respectively.¹⁵

3.2 The Realized Risk Premium is Low

Next, I investigate the relation between subjective risk and realized returns to uncover the “realized risk premium.” Specifically, I regress realized returns on subjective risk,

$$r_{t+h}^i = a^{rea} + \hat{\lambda}^{rea} s_t^i + \epsilon_t^{rea,i}, \quad (1.22)$$

where s_t^i is one of the three subjective risk proxies, r_{t+h}^i is the realized return computed over $h \in \{12, 48\}$ months, a^{rea} is the intercept, $\epsilon_t^{rea,i}$ is a residual, and λ^{rea} is the realized risk premium estimate.

Table II shows that the realized risk premium is positive in four of six specifications but only significant in two. The realized risk premium is, however, much lower than the subjective risk premium. Focusing on column 1, moving from the safest stock ($s_t^{VL, safe} = 0$) to the riskiest ($s_t^{VL, risky} = 1$), increases the four-year realized return from 38% to 48% (8% to 10% annualized) relative to an expected move from 28% to 103%.

Interestingly, $s_t^{VL,i}$ again outperforms the two other proxies in terms of the ability to predict realized returns. It is a significant predictor at both a one- and four-year horizon, and the economic magnitude of the coefficient is relatively high at a one-year horizon. In comparison, β_t^i has a weak relation to realized returns at a one-year horizon and is actually a negative predictor of future four-year returns. $SVIX_t^i$ falls somewhere in between, being a strong forecaster of one-year returns as shown in Martin and Wagner (2019) but a weak forecaster of four-year returns.

¹⁵In Section 8.5, I show the required market return implied by the month-by-month regression in (1.21). The required market return is countercyclical, has a low correlation with the subjective expected market return of retail investors, but a high correlation with the subjective expected market return of professional investors.

Table III. The Realized Risk Premium is Low

	$r_{t,t+48}^i$			$r_{t,t+12}^i$		
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.38 (16.19)	0.46 (17.39)	0.34 (8.67)	0.09 (10.62)	0.12 (12.90)	0.07 (5.76)
$s_t^{VL,i}$	0.10 (2.28)			0.07 (2.55)		
β_t^i		-0.07 (-1.88)			0.01 (0.50)	
SVIX $_t^i$			-0.02 (-0.34)			0.04 (1.45)
Observations	434,927	434,921	81,393	357,921	357,921	68,379
Adjusted R 2	0.001	0.000	0.000	0.001	0.000	0.001

Note: The table shows estimates of the realized risk premium, λ^{rea} , by regressing realized return on subjective risk. The realized return is either over a four-year period ($r_{t,t+48}^i$) or a one-year period ($r_{t,t+12}^i$). The three subjective risk proxies are the safety rank from Value Line ($s_t^{VL,i}$), the market beta estimate from Value Line (β_t^i), and the SVIX $_t^i$ measure from Martin and Wagner (2019). The subjective risk proxies are scaled to lie between 0 and 1 each month, so the coefficients show the change in the dependent variable by moving from the safest (0) to the riskiest (1) firm. The number in parenthesis refers to the t -statistic of the coefficient based on standard errors clustered by firm and quarter. I create realized returns by compounding monthly returns that incorporate delisting returns. If a stock is delisted, I assume that the return for the remaining period is equal to the median across stocks in CRSP.

The Realized and Subjective Risk Premium are Significantly Different

The difference between the realized and subjective risk premium is economically large, but is it statistically significant? In other words, could the difference between $\hat{\lambda}$ and λ^{rea} simply reflect sampling variability? I test this possibility by regressing the difference between the subjective expected return and the subsequent realized return on subjective risk,

$$\tilde{E}_t[r_{t,t+h}^i] - r_{t,t+h}^i = \alpha_0 + \alpha_1 s_t^i + \epsilon_t^i, \quad (1.23)$$

where $\tilde{E}_t[r_{t,t+h}^i]$ is a subjective expected return proxy, $r_{t,t+h}^i$ is the subsequent realized return, s_t^i is a subjective risk proxy, ϵ_t^i is a residual, α_0 is the intercept, and α_1 is the parameter of interest. Table IV shows that α_1 is highly significant across all six specifications with t -statistics ranging from 3 to 13. Therefore, I conclude that sampling variability cannot explain the difference between the subjective and realized risk premium.

Table IV. The Realized and Subjective Risk Premium are Significantly Different

	$\tilde{E}_t^{VL}[r_{t,t+48}^i] - r_{t,t+48}^i$			$\tilde{E}_t^{IBES}[r_{t,t+12}^i] - r_{t,t+12}^i$		
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-0.09 (-3.63)	-0.07 (-2.41)	0.01 (0.16)	-0.02 (-2.41)	-0.02 (-2.14)	0.01 (0.92)
$s_t^{VL,i}$	0.64 (13.60)			0.20 (6.50)		
β_t^i		0.57 (11.71)			0.18 (6.36)	
SVIX $_t^i$			0.56 (7.26)			0.14 (3.59)
Observations	434,927	434,921	81,393	357,921	357,921	68,379
Adjusted R 2	0.02	0.01	0.02	0.01	0.01	0.01

Note: The table estimates the difference between the subjective and realized risk premium, $\lambda - \lambda^{rea}$, by regressing the difference between subjective expected return and realized returns on subjective risk. $\tilde{E}_t^{VL}[r_{t,t+48}^i]$ is the subjective expected return over a 4-year period from Value Line and $r_{t,t+48}^i$ is the subsequent realization, $\tilde{E}_t^{IBES}[r_{t,t+12}^i]$ is the subjective expected return over a 1-year period from I/B/E/S and $r_{t,t+12}^i$ is the subsequent realization. The three subjective risk proxies are the safety rank from Value Line ($s_t^{VL,i}$), the market beta estimate from Value Line (β_t^i), and the SVIX $_t^i$ measure from Martin and Wagner (2019). The subjective risk proxies are scaled to lie between 0 and 1 each month, so the coefficients show the change in the dependent variable by moving from the safest (0) to the riskiest (1) firm. The number in parenthesis shows the t -statistic of the coefficient based on standard errors clustered by firm and quarter. I create realized returns by compounding monthly returns that incorporate delisting returns. If a stock is delisted, I assume that the return for the remaining period is equal to the median across stocks.

4 Subjective Risk and Cash Flow Optimism

In Section 3, I argued that the subjective risk premium is high while the realized risk premium is low. Proposition 1 shows that this pattern is consistent with “risk mispricing,” whereby riskier stocks are more prone to overvaluation. In this section, I show that subjective cash flow expectations are excessively high for risky stocks. Further, several patterns in subjective cash flow expectations are consistent with investors suffering from optimism bias. These results suggest that realized returns of risky stocks are low because irrational cash flow optimism offsets their high required returns.

4.1 Subjective Risk Predicts Cash Flow Forecast Errors

To understand the link between subjective risk and objective mispricing, I regress earnings per share (EPS) forecast errors on subjective risk at the forecast date,

$$\frac{\tilde{E}_t[\text{eps}_{t+12}^i] - \text{eps}_{t+12}^i}{p_t^i} = \gamma_0 + \gamma_1 s_t^i + \epsilon_t^i, \quad (1.24)$$

where $\tilde{E}_t[\text{eps}_{t+12}^i]$ is an EPS forecast over the next fiscal year from either Value Line or I/B/E/S, p_t^i is the stock's price at the beginning of the forecast month, and s_t^i is one of the subjective risk proxies. The key parameter of interest is γ_1 , which shows the relation between subjective risk at time t and the realized forecast errors at time $t + 12$. It has a similar interpretation as γ_1 from the theoretical section, which controlled the relation between subjective risk and objective mispricing. Further, since I scale the subjective risk proxies to lie between 0 and 1, the intercept, γ_0 , shows the average forecast error for the safest stocks ($s_t^i = 0$) and γ_1 shows the average forecast errors for the riskiest ($s_t^i = 1$).

Table V shows that the risk coefficient, γ_1 , is positive, large, and highly statistically significant across all six specifications. Focusing on column 1, the average one-year EPS forecast is 0.20% too *low* for the safest stocks and 1.79% too *high* for the riskiest. Hence, based on this metric, safe stocks are slightly undervalued, while risky stocks are overvalued. Even with relatively few observations, the risk coefficient is almost nine standard errors greater than zero.

Across the subjective risk proxies, $s_t^{VL,i}$ is, again, the strongest predictor of objective mispricing. Among individual stocks, $s_t^{VL,i}$ has the strongest relation with subjective expected returns, realized returns, and objective mispricing. These findings suggest that $s_t^{VL,i}$ could be interesting to analyze separately from its effect on required returns.

Table AIV in the appendix shows that the same results hold for the four-year EPS and four-year price-to-earnings (PE) forecasts. As such, the result that subjective risk predicts cash flow forecast errors holds across different subjective risk proxies ($s_t^{VL,i}$ vs. β_t^i vs. SVIX_t^i), different forecasters (Value Line vs. sell-side analysts from I/B/E/S), different horizons (one vs. four years), and different forecast variables (EPS vs. PE).

The results in this section support the claim that the realized risk premium is lower than the subjective risk premium. In particular, the results in Section 3 relied on the difference between subjective expected returns and the subsequent realization, while this section relies on the difference between subjective cash flow forecast and the subsequent realization. Regardless of the forecast variable, investors seem to overpay for risky stocks based on flawed expectations of their future performance.

Table V. Subjective Risk Predicts Cash Flow Forecast Errors

	Y: $100 \times (\tilde{E}_t^j[\text{eps}_{t+1}^i] - \text{eps}_{t+1}^i)/p_t^i$					
	j: Value Line			j: I/B/E/S		
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-0.20 (-2.69)	0.14 (1.27)	-0.26 (-2.01)	-0.37 (-5.13)	0.24 (2.09)	-0.40 (-2.65)
$s_t^{VL,i}$	1.79 (8.09)			2.62 (8.95)		
β_t^i		1.05 (4.64)			1.31 (5.29)	
SVIX $_t^i$			1.08 (2.91)			1.61 (3.49)
Observations	34,705	34,705	6,329	38,541	38,541	7,282
Adjusted R 2	0.01	0.00	0.01	0.02	0.00	0.01

Note: The table shows estimates from the regression in (1.24) of forecast error on subjective risk. I define the forecast error as the forecast of earnings per share (EPS) in the next fiscal year minus the subsequent realization scaled by the price issued at the beginning of the forecast month. The EPS forecasts are from Value Line in columns 1-3 and I/B/E/S in columns 4-6. For each firm-fiscal year, I only retain the first forecast issued at least 45 days and at most 180 days after the most recent fiscal year’s EPS announcement. The dependent variables are winsorized at the top/bottom 1%. The subjective risk proxies are scaled to lie between 0 and 1 each month, so the intercept shows the forecast error on the safest stocks (risk=0), and the slope coefficient shows the forecast error from the riskiest stock (risk=1). The number in the parenthesis refers to the t -statistic based on standard errors clustered by firm and quarter of the fiscal year-end.

4.2 Optimism Bias Can Explain Risk Mispricing

Why are cash flow forecasts systematically too high for risky stocks? I consider two potential mechanisms: Optimism bias and extrapolative expectations. Optimism bias correctly predicts that average forecast errors are positive, that forecast errors increase in risk, and that forecast errors increase in the forecast horizon. In contrast, extrapolative expectations predict that risky stocks have high past cash flow growth, which is counterfactual.

Plausible Explanation: Optimism Bias

Optimism bias refers to the general tendency for people to be too optimistic about future outcomes. For example, Sharot (2011) reports that a typical estimate is that around 80% of the population are too optimistic while the remaining 20% are too pessimistic. According to Kahneman (2011, p. 255), optimism bias could be the most significant cog-

nitive bias.¹⁶ Cassella et al. (2022) documents optimism bias in GDP and unemployment forecasts, and Cassella et al. (2021) show that optimism bias can explain time-variation in the equity term structure. I focus on the interaction between stock risk and optimism bias.

Proposition 4 shows that optimism bias induces an upwards bias in expectations and that this upwards bias is larger when cash flow uncertainty is higher. Therefore, I test three predictions that follow if forecasters suffer from optimism bias:

1. The average forecast error is positive,
2. the forecast error increases in subjective risk, and
3. the forecast error increases in the forecast horizon.

Prediction 1 follows because optimism bias induces an upwards bias for all stocks, prediction 2 follows because riskier stocks have more uncertain cash flows (see Section 8.2), and prediction 3 follows because cash flow uncertainty is increasing in the forecast horizon.

I test these three predictions jointly by slightly modifying the regression in (1.24). Specifically, to test prediction 1, I regress forecast errors on the demeaned subjective risk such that the intercept captures the average forecast error. In addition, to test prediction 3, I look at multiple forecast horizons. The full specification is,

$$\frac{\tilde{E}_t[\text{eps}_{t+h}^i] - \text{eps}_{t+h}^i}{p_t^i} = \alpha_0^h + \alpha_1^h(s_t^i - \bar{s}_t) + \epsilon_{t+h}^i, \quad (1.25)$$

where h denotes the forecast horizon, \bar{s} is the average subjective risk, and the remaining symbols have the same meaning as in (1.24).

Figure 2 offers support for all three predictions. The left panels show that the intercept, α_0^h , is around zero for EPS forecasts over the next quarter but positive and significant for longer horizons, providing support for prediction 1. For I/B/E/S, the positive forecast errors could reflect incentives-related biases where forecasts are high to please investment banking clients or to generate trading commissions (see, e.g., Kothari (2001)). However, Value Line does not earn money from investment banking or trading commissions, so the results for Value Line provide cleaner evidence of a genuine optimism bias.

The right panels show that the risk coefficient, α_1^h , is positive for all horizons providing support for prediction 2. The effect is highly significant for all horizons except one quarter ahead, and the effect holds for all three subjective risk proxies.

¹⁶Kahneman (2011, p. 255) writes: “The planning fallacy is only one manifestation of a pervasive optimistic bias. Most of us view the world as more benign than it really is, our own attributes as more favorable than they truly are, and the goals we adopt as more achievable than they are likely to be. We also tend to exaggerate our ability to forecast the future, which fosters optimistic overconfidence. In terms of its consequences for decisions making, the optimistic bias may well be the most significant of the cognitive biases.”

Figure 2 also provides support for prediction 3. The intercept and the risk coefficient are both increasing in the forecast horizons, suggesting that any bias in forecasts increases with the forecast horizons. Overall, I conclude that the errors in subjective cash flow expectations are consistent with forecasters suffering from optimism bias.

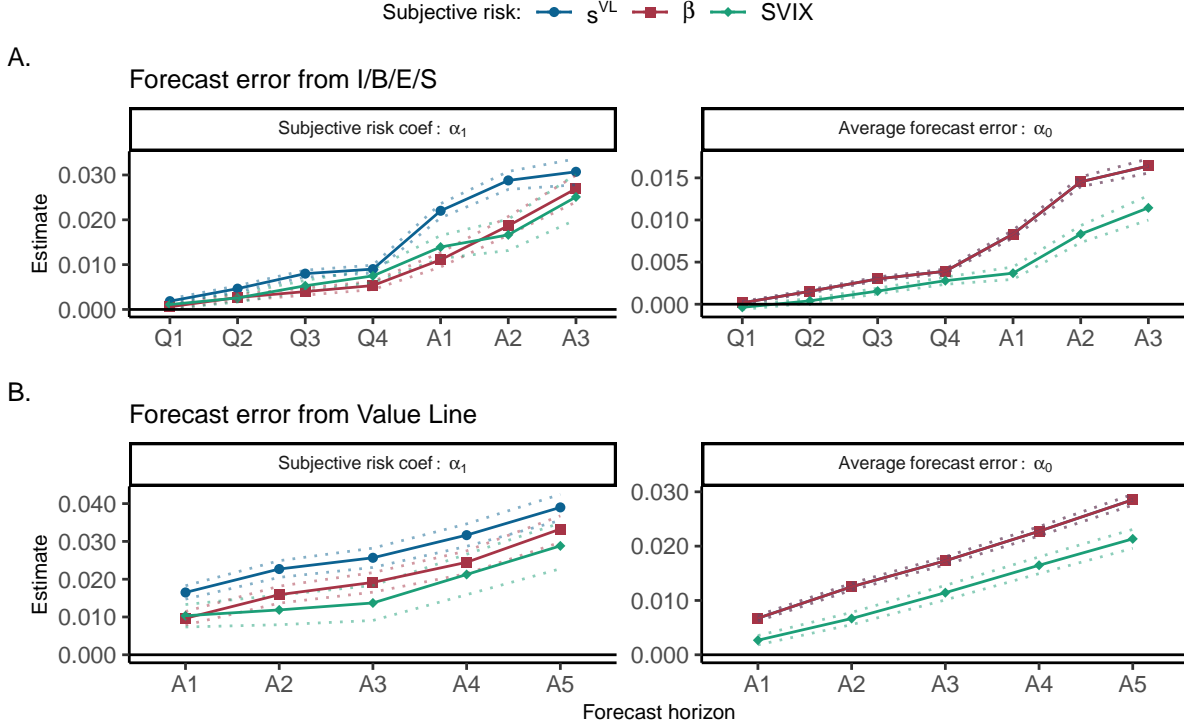


Figure 2. Empirical support for three predictions of optimism bias

Note: The figure shows the α_0^h (left sub-panel) and α_1^h (right sub-panel) coefficients from the regression,

$$\frac{\tilde{E}_t[\text{eps}_{t+h}^i] - \text{eps}_{t+h}^i}{p_t^i} = \alpha_0^h + \alpha_1^h(s_t^i - \bar{s}_t) + \epsilon_{t+h}^i,$$

where $\tilde{E}_t[\text{eps}_{t+h}^i] - \text{eps}_{t+h}^i$ is an earnings per share (EPS) forecast over h months, eps_{t+h}^i is the subsequent realization, p_t^i is the stock's price at the beginning of the forecast month, and $(s_t^i - \bar{s}_t)$ is the demeaned subjective risk of a stock on the forecast date. Optimism bias predicts that α_0^h is positive, α_1^h is positive, and that α_0^h and α_1^h are increasing in h . The EPS forecast is from I/B/E/S in panel A and Value Line in panel B. Subjective risk is either the safety rank from Value Line, s^{VL} , the market beta forecast from Value Line, β , or the SVIX measure from Martin and Wagner (2019). Q1-Q4 means that the forecast horizon is one to four quarters ahead, while A1-A5 means that the forecast is one to five years ahead. The dotted lines show the 95% confidence intervals based on standard errors clustered by stock and the quarter where the fiscal period ends.

Alternative Explanation: Extrapolative Expectations

Figure A4 shows that riskier stocks tend to have higher expected growth in long-term earnings (LTG). La Porta (1996) finds that stocks with high LTG underperform stocks with low LTG because cash flow forecasts are too high LTG stocks. Bordalo et al. (2019)

show that this forecast error can be explained by extrapolative (diagnostic) expectations since high LTG stocks tend to have had high past cash flow growth.

Extrapolative expectations can explain risk mispricing if risky stocks have high past EPS growth. Therefore, I consider the earnings evolution of risky stocks before and after they are classified as “risky,” following the approach in Figure 2 from Bordalo et al. (2019). Specifically, I sort the 10% riskiest and the 10% safest stocks into separate portfolios each month. Within each portfolio-month, I compute the average split-adjusted earnings per share (EPS) in the fiscal year ending in $t - 3, t - 2, \dots, t + 3$, where t is the most recent fiscal year (I winsorize split-adjusted EPS at 1% and 99%), and average these numbers over time.

Figure 3 shows that the EPS of risky stocks generally declines in the years prior to their risk classification but increases in the years after. For stocks classified as risky according to the safety rank from Value Line, the average EPS at the time of portfolio formation is negative, while it is slightly positive three years after. Conversely, safe firms’ EPS increases before and after being classified as “safe.”

The evidence in Figure 3 is inconsistent with extrapolative expectations. With extrapolative expectations, investors would extrapolate the negative EPS trend prior to portfolio formations and thus have overly pessimistic cash flow expectations of risky stocks. In contrast, what I find empirically is overly optimistic cash flow expectations for risky stocks. I, therefore, conclude that extrapolative expectations are unlikely to explain the overvaluation of risky stocks.

5 The Required Return of Equity Factors

I use the subjective risk premium estimates from Table II to compute the required return of 119 equity factors from Jensen et al. (2022a).¹⁷ I then test the “risk hypothesis” that a factor’s average realized return is equal to its required return. I reject this hypothesis for 71% to 79% of the factors (depending on the subjective risk and return proxies used to estimate required returns), suggesting that most factors represent behavioral mispricing rather than rational risk compensation.

Testing the risk hypothesis is crucial for informing theories about why a particular factor works. A rational factor theory assumes that the risk hypothesis being true (see, e.g., Berk et al., 1999; Carlson et al., 2004; Zhang, 2005; Lettau and Wachter, 2007; and Kogan and Papanikolaou, 2013, 2014), while a behavioral factor theory rely on the risk hypothesis being false (see, e.g., Shefrin and Statman, 1985; De Long et al., 1990a,b;

¹⁷Jensen et al. (2022a) studies 153 factors in total, but only 119 have a paper claiming that the factor is a significant predictor of returns.

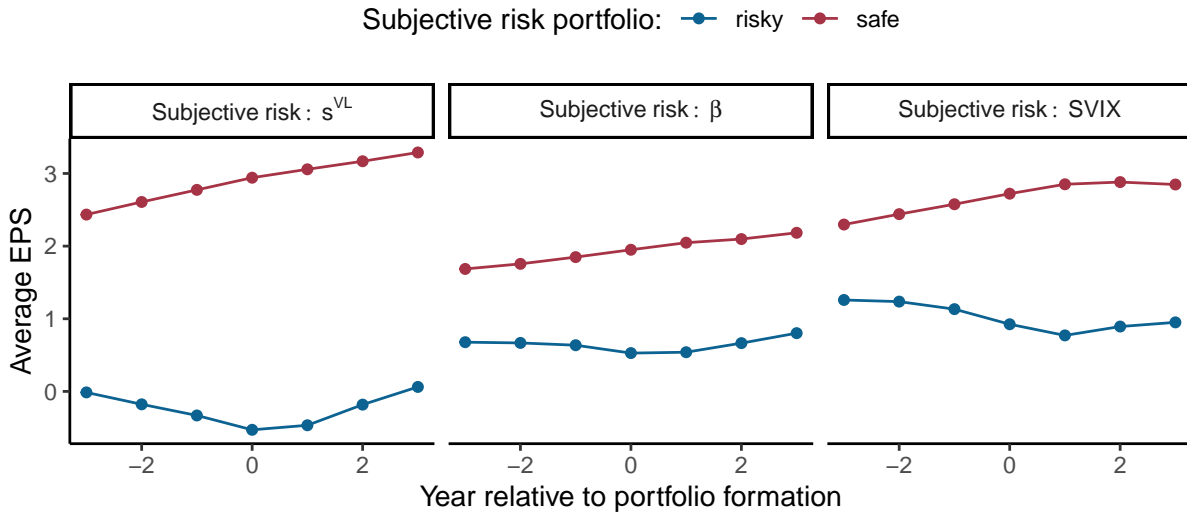


Figure 3. Extrapolative expectations cannot explain risk mispricing

Note: The figure shows the average split-adjusted earnings per share (EPS) of risky and safe stocks. Each month, I sort the 10% riskiest stocks and the 10% safest stocks into separate portfolios, where subjective risk is either the safety rank from Value Line, $s_t^{VL,i}$, the market beta estimate from Value Line, β_t^i , or the risk measure from Martin and Wagner (2019), $SVIX_t^i$. Within each portfolio, I compute the average split-adjusted earnings per share (EPS) in the fiscal year ending in $t-3, t-2, \dots, t+3$, where t is the most recent fiscal year, and I winsorize split-adjusted EPS at 1% and 99%. The points in the figure show these numbers averaged over time.

Daniel et al., 1998; Barberis et al., 1998; Hong and Stein, 1999; Barberis and Huang, 2008; Hong and Sraer, 2016; Bouchaud et al., 2019; and Bordalo et al., 2019).¹⁸

5.1 The Risk Hypothesis is False for Most Factors

To test the risk hypothesis, I need to compute each factor’s realized return and its required return. I follow the procedure in Jensen et al. (2022a) to create a factor’s realized return. I extract all common stocks from CRSP with a non-missing market equity and factor characteristic. Each month, I sort stocks into three portfolios based on the underlying factor characteristic with breakpoints based on non-microcap stocks.¹⁹ Within each portfolio, I weigh stocks using capped value-weights, meaning that a stock’s weight is proportional to its market equity capped at the 80th percentile among NYSE stocks. This weighting scheme avoids an over-reliance on mega-cap stocks while still giving tiny weights to tiny stocks. The realized factor return is the return of the long portfolio minus that of the short portfolio. Table AI contains the names of the 119 factor characteristics and whether the corresponding factor goes long high or low values of the characteristic.

¹⁸An interesting hybrid is Barberis et al. (2021) who tests whether prospect theory can explain factors.

¹⁹Microcap stocks have market equity below the 20th percentile of NYSE stocks (Fama and French, 2008).

I compute a factor's required return following a similar approach,, but I am constrained to stocks with a non-missing subjective risk. Since these stocks tend to be large, I use breakpoints from all stocks such that the long and short portfolio contains the same number of stocks.

I start by computing the required return on the three portfolios $j \in \{\text{long, middle, short}\}$. To do so, I compute the weighted subjective risk of stocks in the portfolio each month, and then take an average over time to get the unconditional portfolio risk,

$$\bar{s}_t^j = \frac{1}{T} \sum_{t=1}^T \sum_{i \in \{j_t\}} \pi_t^{i,j} s_t^i, \quad (1.26)$$

where T is the total number of months, $\{j_t\}$ is the stocks in portfolio j in month t , $\pi_t^{i,j}$ is the portfolio weight of stock i , and s_t^i is the stock's subjective risk.

I then use the subjective risk premium estimates from Table II to convert the subjective portfolio risk into a required return. The required returns based on these estimates are over a $h = 12$ or $h = 48$ month horizon (depending on the subjective expected return proxy), so I convert them to a monthly horizon,²⁰

$$\bar{r}_{t+1}^{req,j} = \left(\bar{r}_{t,t+h}^f + a + \hat{\lambda} \times \bar{s}_t^j \right)^{1/h} - 1 - \bar{r}_{t,t+1}^f, \quad (1.27)$$

where $\bar{r}_{t+1}^{req,j}$ is the monthly required return on the portfolio, a and $\hat{\lambda}$ are the parameters from Table II, h is the horizon of the subjective expected return proxy, and $\bar{r}_{t,t+h}^f$ is the average risk-free rate over h months during the sample period used to estimate the parameters.

The required return on the factor is the required return on the long portfolio minus the required return on the short portfolio:

$$\bar{f}_{t+1}^{req} = \bar{r}_{t+1}^{req,long} - \bar{r}_{t+1}^{req,short}. \quad (1.28)$$

I test the risk hypothesis using following test statistic,

$$\tau = \bar{f}_{t+1} - \bar{f}_{t+1}^{req}, \quad (1.29)$$

where \bar{f}_{t+1} is the average realized return on the factor. The null hypothesis (i.e., the risk

²⁰The required returns from Table II are excess returns, so to convert it into a monthly horizon, I first add the risk-free rate over the original horizon and then subtract the one-month risk-free rate: $\bar{r}_{t,t+1}^{req} = (1 + r_{t,t+h}^f + r_{t,t+h}^{req})^{1/h} - 1 - r_{t,t+1}^f$, where $r_{t,t+h}^{req}$ is the required return over h months and $r_{t,t+h}^f$ is the risk-free rate over h months.

hypothesis) is that the factor's average realized return is equal to its required return,

$$H_0^{risk} : \bar{f}_{t+1} = \bar{f}_{t+1}^{req}. \quad (1.30)$$

What remains is to derive the distribution of the test statistic under the null hypothesis.

Theoretical Null Distribution

I have to navigate a range of issues in testing the null hypothesis. First, the required return data covers a shorter period than the realized return data. I could limit the sample to the overlapping period, but doing so would make the expected returns estimate less precise. Instead, I incorporate information from the non-overlapping and overlapping period by considering the weighted factor return, $\bar{f}_{t+1} = w_1 \bar{f}_{t+1}^1 + (1 - w_1) \bar{f}_{t+1}^2$, where $w_1 = T_1 / (T_1 + T_2)$ is the number of months in the non-overlapping period divided by the total number of observations and \bar{f}_{t+1}^1 and \bar{f}_{t+1}^2 is the average factor return in each of the two periods. I assume that the sampling noise of \bar{f}_{t+1}^1 is uncorrelated with that of \bar{f}_{t+1}^2 and \bar{f}_{t+1}^{req} . In contrast, I allow for correlation between \bar{f}_{t+1}^2 and \bar{f}_{t+1}^{req} . Finally, I assume that \bar{f}_{t+1}^1 , \bar{f}_{t+1}^2 , and \bar{f}_{t+1}^{req} are multivariate normally distributed, $x \sim N(\mu, \Sigma)$, where

$$x = \begin{bmatrix} \bar{f}_{t+1}^1 \\ \bar{f}_{t+1}^2 \\ \bar{f}_{t+1}^{req} \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & \text{Cov}(\bar{f}_{t+1}^2, \bar{f}_{t+1}^{req}) \\ 0 & \text{Cov}(\bar{f}_{t+1}^2, \bar{f}_{t+1}^{req}) & \sigma_3^2 \end{bmatrix}. \quad (1.31)$$

Letting $w = [w_1, 1 - w_1, -1]'$, the distribution of the test statistic, τ from (1.29), under the null is

$$\tau | H_0^{risk} \sim N(0, w' \Sigma w), \quad (1.32)$$

where $w' \Sigma w = w_1^2 \sigma_1^2 + (1 - w_1)^2 \sigma_2^2 + \sigma_3^2 - 2(1 - w_1) \text{Cov}(\bar{f}_{t+1}^2, \bar{f}_{t+1}^{req})$. The p -value (the two-sided probability of the observed test statistic under the null) is therefore,

$$p = 2\Phi\left(\frac{-|\tau|}{\sqrt{w' \Sigma w}}\right), \quad (1.33)$$

where Φ is the cumulative distribution function of a standard normal variable. I reject the risk hypothesis if the p -value is below 5%.

Estimation

Next, I need to account for two additional issues in estimating the null distribution. First, the required return is estimated rather than a fixed function of subjective risk, which adds estimation noise to σ_3^2 and $\text{Cov}(\bar{f}_{t+1}^2, \bar{f}_{t+1}^{req})$. Second, required returns are persistent.

I therefore estimate the parameters in the overlapping period, σ_2^2 , σ_3^2 , and $\text{Cov}(\bar{f}_2, \bar{f}_{t+1}^{req})$, via a moving block bootstrap procedure. Each block contains all data from a range of temporal months, but the specific block size depends on the subjective expected return proxy. For Value Line the block size is $n = 30$ months while it is $n = 12$ for I/B/E/S. In each bootstrap iteration, I create a bootstrap sample by randomly choosing $\lceil T_2/n \rceil$ blocks with replacement. I then delete the last $\lceil T_2/n \rceil n - T_2$ months from the last block to ensure that the bootstrap sample covers the same number of months as the original sample.

Within each bootstrap sample, I compute the average realized factor return, $\bar{f}_{t+1}^{2,b}$, the average subjective risk of the long, $\bar{s}_t^{long,b}$, and short $\bar{s}_t^{short,b}$ portfolio. Furthermore, I re-estimate the regression from Table II that maps subjective risk to required returns to account for estimation noise. Using these inputs in (1.27) and (1.28), I estimate the factor's required return, $\bar{f}_{t+1}^{req,b}$. I repeat this procedure 1,000 times.

The covariance matrix of these bootstrap realization gives me an estimate of σ_2^2, σ_3^2 and $\text{Cov}(\bar{f}_{t+1}^2, \bar{f}_{t+1}^{req})$. I complete Σ by estimating the sampling variation of \bar{f}_{t+1}^1 as $\sigma_1^2 = \text{Var}(f_{t+1}^1)/T_1$, where $\text{Var}(f_{t+1}^1)$ is the factor's realized return variance in the non-overlapping period.

Results

Figure 4 shows the components and results of the risk hypothesis test.²¹ The y -coordinate shows each factor's average realized return and is the same in all panels. The x coordinate shows the factor's required return implied by a specific combination of subjective risk and return expectations. For example, in the top-left panel, required returns are based on the safety rank and return expectation from Value Line. To make factors easier to compare, I scale their realized and required returns to an ex-post volatility of 10%.²² The figure also includes a 45° line, representing the risk hypothesis that the realized return equals the required return. A point far from the dotted line provides evidence against the risk hypothesis. Red points show factors where I reject the risk hypothesis (the p -value from (1.33) is below 5%).

I reject the risk hypothesis for 73.9% of the factors when the required return is based on the safety rank and expected return from Value Line (top left panel). This number suggests that the realized returns of most factors do not only reflect rational compensation for risk. This conclusion is robust across the six variations of required returns, with risk rejection rates ranging from 71.4% to 79.0%.

²¹Table AII shows the detailed results for each factor.

²²For example, for a factor with a realized return volatility of 20%, I multiply its realized and required return by 0.5. This standardization is only for visualization purposes and does not affect the risk hypothesis test.

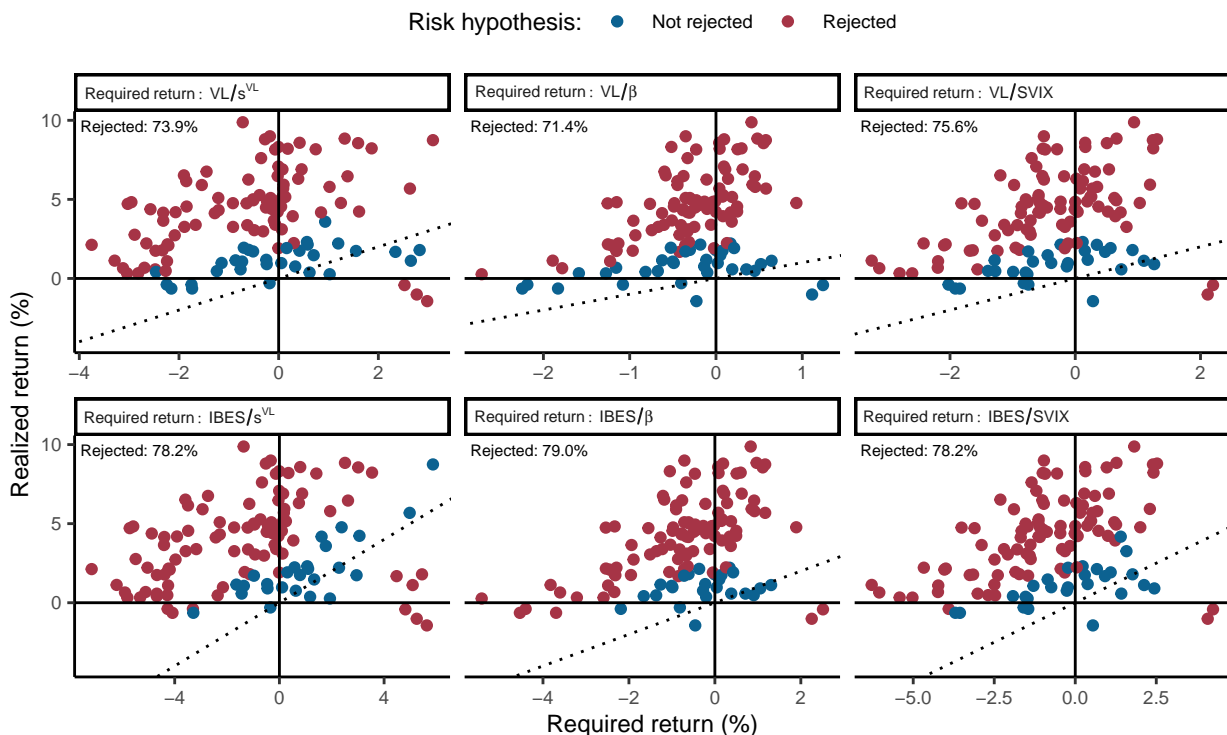


Figure 4. The Risk Hypothesis is False for Most Equity Factors

Note: The figure shows the required return and the realized return of 119 equity factor for six different required return proxies. The top-left corner shows, on the x -axis, the required return of the 119 factors, based on the safety rank and subjective expected returns from Value Line, while the y -axis shows the average realized return of each factor. The remaining panels differ only in the subjective risk and return proxies used to compute the factors' required returns, which are shown in the panel titles. The name before the slash refers to the subjective expected return proxy which is either from Value Line or I/B/E/S (IBES). The name after the slash refers to the subjective risk proxy where s^{VL} is the safety rank from Value Line, β is the market beta estimate from Value Line, and SVIX is the risk measure from Martin and Wagner (2019). The solid line shows the 45° line, which corresponds to equality between the required return and the realized return. The coloring shows whether the p -value from (1.33) is below 5%; that is, whether the risk hypothesis is rejected. The number in the top-left corner of each panel shows the fraction of factors where the risk hypothesis is rejected. Table AII presents the detailed information about the realized and required return for each factor.

One reason for the high rejection rate is that many factors have a positive realized return despite a negative required return. This finding is challenging to reconcile with a rational explanation since a negative required return implies that stocks in the long portfolio are subjectively safer than stocks in the short portfolio. To test the severity of this issue, I change all factors to have a positive realized return and then test whether the required return is significantly negative. To do so, I compute the p -value of the required return (the probability of the observed required return if the true required return was zero) as,

$$p^{req} = \Phi \left(\frac{\bar{f}_{t+1}^{req}}{\sigma_3} \right), \quad (1.34)$$

where \bar{f}_{t+1}^{req} is the factor’s required return, σ_3 is the standard error of the required return, and $\Phi(x)$ is the standard normal cumulative distribution function. I categorize the required returns as significantly negative if p^{req} is below 5%, significantly positive if p^{req} is above 95%, and insignificant otherwise.

Table VI shows that 44%-50% of the factors have a significantly negative required returns, while only 14%-27% have a significantly positive required return. Disregarding statistical significance, the fraction of factors with a negative required returns is 61%-63%. Overall, most factors bet on safe, not risky, stocks.

Table VI. Many Factors have a Negative Required Return

Required return proxy		Negative	Insignificant	Positive
<i>VL</i>	/ $s_t^{VL,i}$	0.47	0.26	0.27
<i>VL</i>	/ β_t^i	0.44	0.42	0.14
<i>VL</i>	/ $SVIX_t^i$	0.50	0.27	0.24
<i>IBES</i>	/ $s_t^{VL,i}$	0.49	0.24	0.27
<i>IBES</i>	/ β_t^i	0.50	0.34	0.16
<i>IBES</i>	/ $SVIX_t^i$	0.50	0.27	0.23

Note: The table shows the fraction of factors with a significant required return. To compute these fractions, I first change the direction of all factors such that the long portfolio has a higher realized return than the short portfolio. I then calculate the p^{req} value from (1.34) on the modified portfolio. The “Negative” columns show the fraction with p^{req} below 5%, “Positive” shows the fraction with p^{req} above 95%, and “Insignificant” shows the remaining fraction. The subjective expected return either comes from Value Line (VL) or I/B/E/S (IBES). The subjective risk is either the safety rank from Value Line ($s_t^{VL,i}$), the market beta from Value Line (β_t^i), or the $SVIX_t^i$ measure from Martin and Wagner (2019).

Figure 5 highlights the relation between realized and required returns for 13 prominent equity factors. The figure also shows how the factor’s required return varies across the subjective risk proxies (all three points use the expected return from Value Line). The three risk proxies largely agree on the result of the risk hypothesis test, but there is substantial disagreement about the required return of some factors. For example, the size factor that overweight small stocks is unrelated to risk when risk is measured using a stock’s market beta. In contrast, according to Value Line’s safety rank, small stocks are much riskier than larger stocks. Another example is the distress factor, which shorts distressed stocks and buys stable stocks. Distressed and stable stocks have comparable market betas, but distressed stocks have a much higher safety rank. Small and distressed stocks are risky according to Value Line but not according to the CAPM.

The three risk proxies also have points of agreement: The low beta factor overweight safe stocks, and the momentum factor is unrelated to risk.

Irrespective of the specific risk proxy, some factors, such as the quality, low equity issuance, and profitability factor, have a large and positive return despite stocks in the

long portfolio being much safer than those in the short portfolio. Conversely, the factor that overweight young stocks have a negative realized return despite young stocks being relatively risky. Other factors such as the momentum, book-to-market, duration, and asset growth factor have a weak relation to subjective risk.

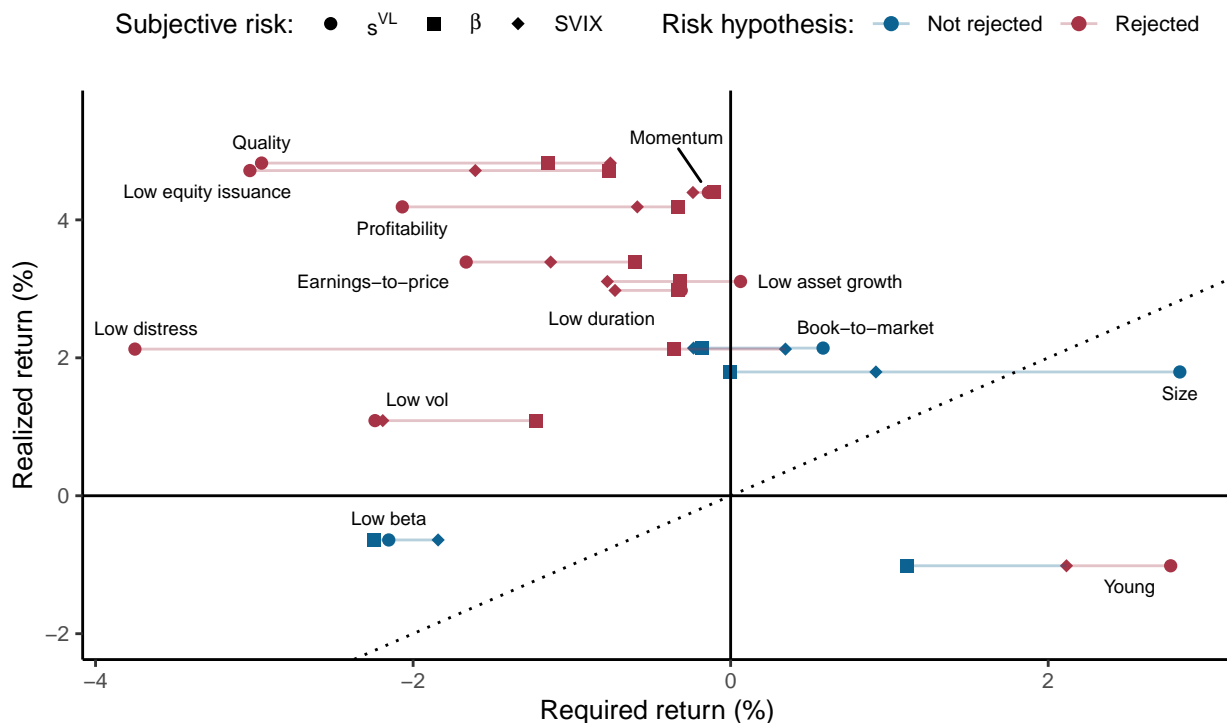


Figure 5. The Realized and Required Return of Prominent Equity Factors

Note: The y coordinate shows each factor’s average realized return, and the x coordinate shows the factor’s required return across different proxies. The circle uses the safety rank from Value Line ($s_t^{VL,i}$), the square uses the market beta from Value Line (β_t^i), and the diamond uses the $SVIX_t^i$ measure from Martin and Wagner (2019). All three proxies use the subjective expected return from Value Line. Blue points indicate that a factor’s realized return is significantly different from the factor’s required return, while red points indicate the opposite. The factor name starts with “low” if the factor overweight stocks with a low value of the factor characteristic. The factors, with factors characteristic names from Table AI show in parenthesis, are: Low asset growth (`at_gr1`), low beta (`beta_60m`), book-to-market (`be_me`), low distress (`o_score`), low duration (`eq_dur`), earnings-to-price (`ni_me`), low equity issuance (`chcsho_12m`), momentum (`ret_12.1`), profitability (`ope_be`), quality (`qmj`), (low) size (`market_equity`), low vol (`rvo1_21d`), and young (`age`).

5.2 Risky Stocks are Small, Volatile, and Distressed

The analysis in the previous section, reveal the univariate relation between a stock characteristic and required returns. In this section, I analyze the multivariate drivers of required returns. Required returns are the product of a constant subjective risk premium and subjective risk (1.3), so all variation in required returns is driven by subjective risk. I, therefore, inspect the drivers of subjective risk directly.

I regress subjective risk on eight prominent stock characteristics,²³

$$s_t^i = \alpha_0 + \sum_{k=1}^8 \alpha_k x_t^{k,i} + \epsilon_t^i, \quad (1.35)$$

where $x_t^{k,i}$ is the value of the k^{th} characteristic for a particular stock. I standardize the characteristics by using cross-sectional percentile ranks, which means that the magnitude of the coefficients are informative about their relative importance.²⁴ The dependent variable is either the safety rank from Value Line, $s_t^{VL,i}$, or the SVIX measure from Martin and Wagner (2019). I do not include the market beta estimate from Value Line because this measure is, by assumption, driven solely by a stock's market beta.

The top panel in Figure 6 shows that, according to the safety rank, a stock is risky if it is small, has a high return volatility, a high market beta, and a low distance to default. Said differently, a risky stock is small, volatile, and distressed. Furthermore, risky stocks have a high expected long-term earnings growth as shown in Section 4.2.

The bottom panel in Figure 6 shows that, according to the SVIX measure, high risk primarily reflects high return volatility. The SVIX measure reflects a stock's risk-neutral variance, so it is, perhaps, unsurprising that this measure correlates highly with a stock's real-world volatility.

Asset growth, book-to-market, and profitability have a limited marginal impact on both measures. Hence, the risk of stocks sorted by these characteristics reflects their implicit exposure to other characteristics. For example, profitable stocks most likely have a high distance to default, which makes them subjectively safe.

Return volatility and distance to default, both important drivers of subjective risk, are sometimes thought to capture idiosyncratic, rather than systematic, risk. To test whether subjective risk captures systematic or idiosyncratic risk, I rely on the fact that idiosyncratic risk disappears in well-diversified portfolios (Markowitz, 1952). To test whether subjective risk captures systematic risk, I sort stocks into three portfolios according to each risk proxy and compute the return on the risky-minus-safe portfolio. The risky-minus safe portfolios have a high return volatility (17% for the safety rank, 17% for market beta, 21% for SVIX), a high market beta (0.72 for the safety rank, 0.86 for market beta, 0.96 for SVIX), and a high consumption beta (1.02 for the safety rank, 1.07 for market beta, and 1.51 for SVIX). These statistics suggest that stocks with a high subjective risk have a high systematic risk.

²³The eight characteristics are the market beta estimate from Value Line, the long-term growth rate in EPS forecast from Value Line, asset growth, book-to-market, distance to default, and market equity, profitability, and return volatility.

²⁴I create a cross-sectional percentile rank by ranking stocks according to a characteristic within a month and then dividing by the number of stocks. This procedure handles outliers and ensures that all characteristics lie on a common scale between 0 and 1.

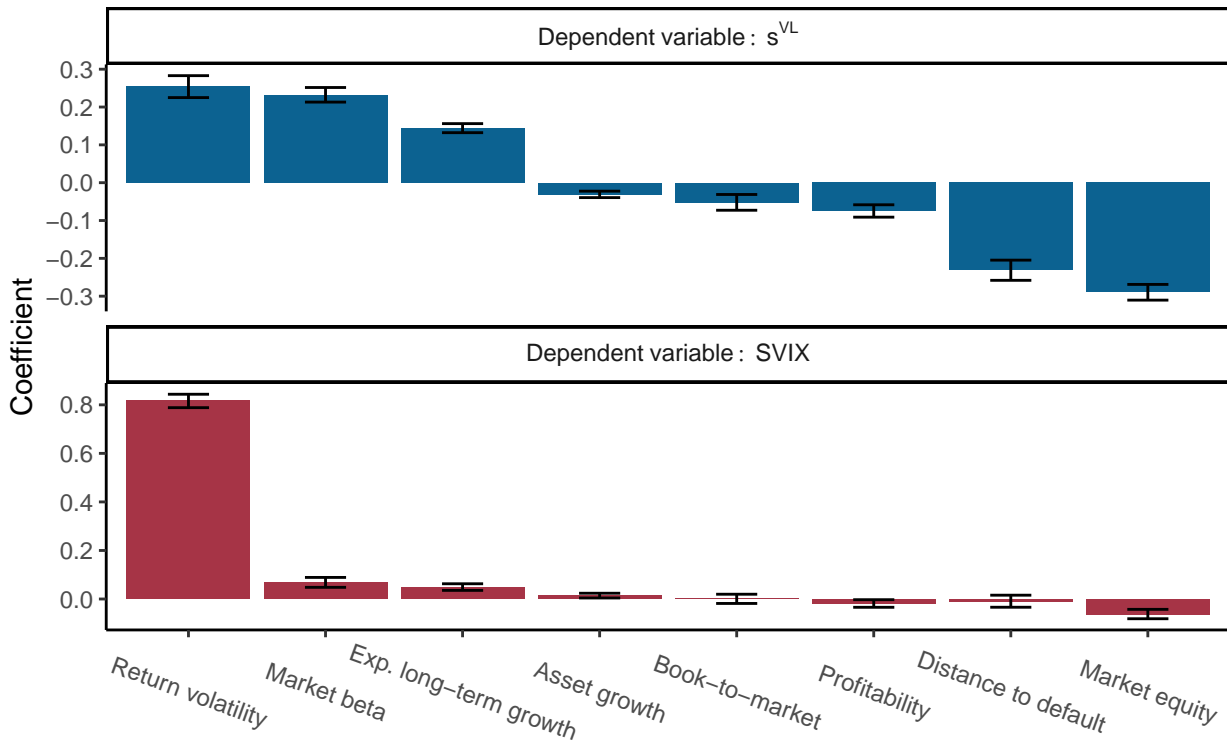


Figure 6. The Drivers of Required Returns

Note: The figure shows the coefficient estimates and the corresponding 95% confidence interval from regressing subjective risk on eight stock characteristics as in (1.35). The subjective risk proxy is either the safety rank, s^{VL} , from Value Line (top panel) or the risk measure, SVIX, from Martin and Wagner (2019) (bottom panel). The adjusted R^2 of the regressions are 70% in the top panel and 84% in the bottom panel. Return volatility is the standard deviation of daily returns over the past 252 trading days, market beta is the market beta estimate from Value Line, exp. long-term growth is Value Line’s estimate of EPS growth over the next three-to-five years, asset growth is the percentage change in total assets over the past fiscal year, book-to-market is the book-to-market equity ratio, profitability is computed as in Fama and French (2015), distance to default is computed using the method in Bharath and Shumway (2008), and market equity is the market equity value of the security.

6 Can Asset Pricing Models Explain Required Returns?

Asset pricing models have a dual mandate to (1) predict realized returns and (2) explain required returns. Nevertheless, most asset pricing models are judged solely by their performance on the “realized return” mandate (see, e.g., Fama and French, 1993a; Fama and French, 2015; Hou et al., 2015; and Barillas and Shanken, 2018). I want to understand whether the best models for the realized return mandate are also the best for the “required return” mandate.

Proposition 3 showed that the answer is “yes” if the market is efficient. In an efficient market, the predictable part of a stock’s realized returns only reflects its required return, so a model that explains realized returns also explains required returns. In an inefficient

market, realized returns also reflect mispricing, which the best model of realized returns must target. By targeting this mispricing, the realized return model sacrifices its ability to explain required returns meaning that the best model for the two mandates differs.

I find that recent empirical models, such as the Fama-French five-factor model, explain realized returns well but required returns poorly. In contrast, traditional models, especially the CAPM, explain required returns well but realized returns poorly. This finding suggests that the search for better models of realized returns has resulted in models that capture mispricing rather than required returns.

6.1 Candidate Asset Pricing Models

I consider nine candidate asset pricing models shown in Table VII, which I divide into three “traditional,” four “recent,” and two “benchmark” models. The traditional models were published before 1994, and two of three are theoretically motivated. In contrast, the recent models were published after 2014 and are predominantly empirically motivated. Finally, the two benchmark models are new to this paper.

Table VII. Candidate Asset Pricing Models

Name	Abbreviation	Reference	Type
CAPM	CAPM	Sharpe (1964), Lintner (1965), Mossin (1966)	Traditional
Consumption CAPM	CCAPM	Breeden (1979)	Traditional
Fama-French-3	FF3	Fama and French (1993a)	Traditional
Fama-French-5	FF5	Fama and French (2015)	Recent
Investment CAPM	HXZ	Hou et al. (2015)	Recent
Mispricing Factors	SY	Stambaugh and Yuan (2017)	Recent
Behavioral Factors	DHS	Daniel et al. (2020)	Recent
Machine Learning	ML	This paper	Benchmark
Required Return Model	REQ	This paper	Benchmark

Note: The table shows information about the nine candidate asset pricing models I evaluate in Section 6. The “name” and “abbreviation” columns show the naming and abbreviation convention I use to refer to the models throughout the paper, “reference” shows the paper(s) that proposed the model, and “type” shows my classification scheme. The two benchmark models are new to this paper.

I test the candidate models on the $3 \times 119 = 357$ high-middle-low portfolios that underlie the factors from Section 5. For each portfolio, I compute its expected return implied by each model. I then test how well these model-implied expected returns align with the portfolio’s average realized and required returns, respectively. A model that explains realized (required) return well assigns high expected returns to portfolios with a high average realized (required) return.

I start by explaining how I compute the expected return implied by each model. I use realized return from 1972 to 2021 to estimate factor loadings and factor premiums because all portfolios and all models except “REQ”, have data available throughout this period.

The first traditional model I consider is the CAPM of Sharpe (1964), Lintner (1965), and Mossin (1966). To compute the CAPM-implied expected return of a portfolio j , I regress the realized return of the portfolio on the contemporaneous market return from 1972 to 2021, $r_t^j = \alpha^j + \beta^j r_t^{mkt} + \epsilon_t^j$. The implied expected return then follows from removing the intercept (which the model implies is zero),

$$E^{\text{CAPM}}[r_{t+1}^j] = \beta^j E[r_{t+1}^{mkt}], \quad (1.36)$$

where $E[r^{mkt}]$ is the expected market return, estimated as the average realized market return from 1972 to 2021.

The second traditional model is the consumption CAPM of (Breedon, 1979, CCAPM). I use a linear approximation based on lognormal consumption and an investor with time-separable power utility (Campbell, 2017, p. 162-163). For each portfolio, I compute the monthly covariance between its realized return and log changes in consumption per capita from 1972 to 2021, $\text{Cov}(r_t^j, \Delta c_t^j)$. The implied expected return is then,

$$E^{\text{CCAPM}}[r_{t+1}^j] = \gamma \text{Cov}(r_t^j, \Delta c_t^j), \quad (1.37)$$

where γ measures the investor's relative risk aversion. I use $\gamma = 10$, which is at the high end of reasonable relative risk aversions (Mehra and Prescott, 1985).

The third traditional model is the three-factor model from Fama and French (1993a, FF3), which adds a value and size factor to the CAPM. This model is the first of five multifactor models, so I explain how I obtain the expected return from a generic K factor model. First, I regress the factor return on the contemporaneous return of the model's pricing factors from 1972 to 2021, $r_t^j = \alpha^j + \sum_{k=1}^K \beta^{j,k} r_t^k + \epsilon_t^j$. For a multifactor model m , the implied expected return is then,

$$E^m[r_{t+1}^j] = \sum_{k=1}^K \beta^{j,k} E[r_{t+1}^k], \quad (1.38)$$

where $E[r^k]$ is the expected return on the k^{th} pricing factor estimated as this factor's average realized return from 1972 to 2021.

The model-implied expected return of the next four models uses (1.38). The first recent model is the five-factor model of Fama and French (2015, FF5). This model is motivated by the dividend discount model and adds a profitability and investment factor to FF3. The next model is the investment CAPM from Hou et al. (2015, HXZ). This model is motivated by the q theory of investments and uses a market, size, profitability, and investment factor.

So far, the models I have considered have a rational justification for "why" they

work. By contrast, the two remaining models have a behavioral justification. The model from Stambaugh and Yuan (2017, SY) is motivated by persistent mispricing, especially overvaluation, which rational investors fail to correct due to arbitrage asymmetry. The model uses a market factor, a size factor, and two mispricing factors related to managerial actions and firm performance. The model from Daniel et al. (2020, DHS) is designed to capture short- and long-horizon mispricing due to investor biases. The model uses a market factor and two behavioral factors. The first behavioral factor is based on the post-earnings announcement drift and captures short-run mispricing. The second behavioral factor is based on equity issuance and captures long-run mispricing.

The two benchmark models are meant to reflect the optimal model for realized returns (ML) and the optimal model for required returns (REQ). ML is based on machine learning forecasts of realized returns at the stock level using an approach similar to Gu et al. (2020a) and Jensen et al. (2022). I predict the returns of individual stocks using a gradient boosted decision tree methodology called “XGBoost” from Chen and Guestrin (2016). For each stock, the inputs to the model is the 119 equity factor characteristics from Section 5 and the outcome variable in the excess return one month ahead. I update the model each decade such that all predictions are out-of-sample, and the first out-of-sample prediction is in 1972. Section 8.8 describes the procedure in detail. The expected return of a portfolio follows from a bottom-up aggregation of the stock level forecasts, and the unconditional expected return of a factor is the average of these predictions from 1972 to 2021,

$$E^{ML}[r_{t+1}^j] = \frac{1}{T} \sum_{t=1}^T \sum_{i \in \{j_t\}} \pi_t^{i,j} \hat{f}_t(x_t^i), \quad (1.39)$$

where $\{j_t\}$ is the stocks in portfolio j at time t , T is the number of months in the sample, $\pi_t^{i,j}$ is a stock’s weight in the portfolio, $\hat{f}(x)$ is the prediction from the estimated XGBoost model, and x_t^i a vector with the 119 stock characteristics.

Finally, the second benchmark model is a required return model (REQ). This model sets a portfolio’s expected return equal to its required return computed as in Section 5,

$$E^{\text{REQ}}[r_{t+1}^j] = \left(\bar{r}_{t,t+h}^f + a + \hat{\lambda} \times \bar{s}_t^j \right)^{1/h} - 1 - \bar{r}_{t,t+1}^f, \quad (1.40)$$

where \bar{s}_t^j is the average subjective risk of the portfolio, a and $\hat{\lambda}$ is the parameters from Table II, h is the horizon of the subjective expected return proxy, and $\bar{r}_{t,t+h}^f$ is the average risk-free rate over a horizon of h for the period used to estimate the parameters. There are six different versions of REQ depending on the subjective risk and return proxies used to estimate the parameters.

To evaluate the models’ ability to explain realized returns, I compute each portfolio’s

average realized return from 1972 to 2021. To evaluate the models' ability to explain required returns, I compute each portfolio's required return using the approach in Section 5.

The level of required returns is generally higher than the level of realized returns (as seen in Figure 1). I, therefore, demean model-implied expected returns, realized returns, and required returns to focus on the model's cross-sectional pricing ability.

6.2 Asset Pricing Models Explain Realized *or* Required Returns

Figure 7 shows the extent to which the model-implied expected returns align with the average realized and required return of the 357 portfolios. Each panel differs only on the x coordinate, which shows a portfolio's model-implied expected return. The y -coordinate of the red triangles shows the corresponding portfolio's average realized return. If a model perfectly explains realized returns, all the red triangles would lie on the dotted 45° line. The required return model and the traditional models explain realized returns poorly. In fact, they tend to imply high expected returns to portfolios with low realized returns. The consumption CAPM is an exception, but it tends to imply a low dispersion in expected returns—another manifestation of the equity premium puzzle (Mehra and Prescott, 1985). In contrast, the recent empirical models explain realized returns well since they tend to imply a high expected return to portfolios with high realized returns. They do, however, seem to be dominated by the ML model.

The y -coordinate of the blue circles is the required return of the portfolios based on the safety rank and expected return from Value Line. The required return model perfectly explains required returns by assumption. More interestingly, required returns closely align with the expected return from traditional models like the CAPM and FF3. In contrast, the expected return of the recent models and the ML model is negatively related to required returns.

To capture the intuition from Figure 7, I create two metrics that summarize each model's ability to explain realized and required returns. I define the “realized pricing ability” of a model as the R^2 from the following model, $\bar{r}_{t+1}^j = a + E^m[r_{t+1}^j] + \epsilon^j$, where \bar{r}^j is the average realized return of each portfolio, $E^m[r_{t+1}^j]$ is the model-implied expected return and a is an intercept that captures level-differences. Concretely, a model's realized pricing ability is,

$$R_m^2 = 1 - \frac{\sum_{j=1}^{357} [(E^m[r_{t+1}^j] - c_1) - (\bar{r}^j - c_2)]^2}{\sum_{j=1}^{357} (\bar{r}^j - c_2)^2}, \quad (1.41)$$

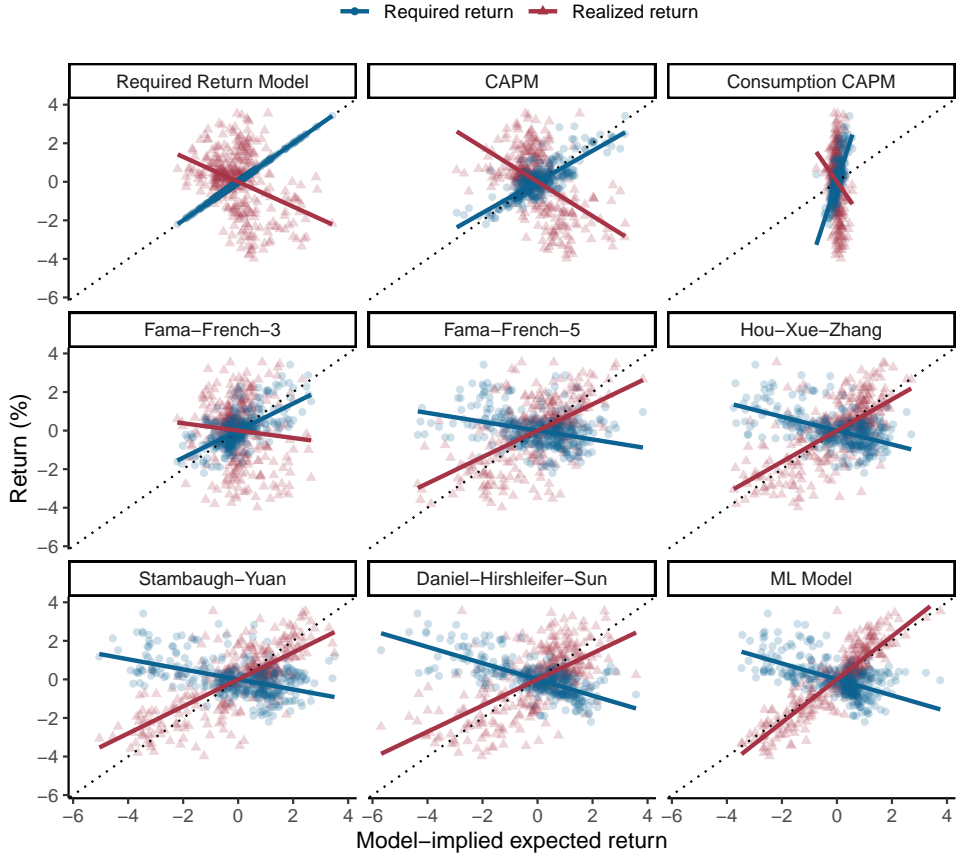


Figure 7. Model-Implied Expected Return vs. Required and Realized Returns

Note: The table visualizes the extent to which model-implied expected returns align with realized and required returns. The test assets are the 357 high-middle-low portfolios underlying the equity factors described in Table AI. For each model-portfolio pair, the x-coordinate is the model-implied expected return of the portfolio, while the y-coordinate is the portfolio's required return (blue circle) or its average realized return (red triangle). The y and x variables have been demeaned. The required returns is based on $\tilde{E}_t^{VL}[r_{t,t+48}^i]$ and $s_t^{VL,i}$. The blue (red) solid line is the best linear fit between model-implied expected returns and required returns (average realized returns). The dotted line is the 45° line that indicates a perfect match between the x and y variable. The models are described in Table VII.

where $c_1 = \frac{1}{357} \sum_{j=1}^{357} E^m[r_{t+1}^j]$ is the average model-implied expected return across the 357 portfolios, and $c_2 = \frac{1}{357} \sum_{j=1}^{357} \bar{r}_{t+1}^j$ is the average realized return. Similarly, a model's required pricing ability is the R^2 from the model, $\bar{r}_{t+1}^{req,j} = \tilde{a} + E^m[r_{t+1}^j] + \tilde{\epsilon}^j$, which is,

$$\tilde{R}_m^2 = 1 - \frac{\sum_{j=1}^{357} [(E^m[r_{t+1}^j] - c_1) - (\bar{r}_{t+1}^{req,j} - c_3)]^2}{\sum_{j=1}^{357} (\bar{r}_{t+1}^{req,j} - c_3)^2}, \quad (1.42)$$

where $\bar{r}_{t+1}^{req,j}$ is the portfolio's required return and $c_3 = \frac{1}{357} \sum_{j=1}^{357} \bar{r}_{t+1}^{req,j}$ is the average required return across portfolios.

The first row of Table VIII presents the realized pricing ability of the nine models.

Consistent with Figure 7, the traditional models perform poorly with R^2 ranging from -0.92 for CAPM to -0.05 for CCAPM. In contrast, the recent models perform well with R^2 ranging from 0.26 for FF5 to 0.45 for SY4. According to this metric, the recent empirical models are superior to the traditional model for explaining realized returns. Still, the ML model is superior to the recent model with an R^2 of 0.78. This substantial outperformance is surprising, considering that the ML model makes out-of-sample predictions while the asset pricing models rely on in-sample factor loadings and factor premiums. This result suggests that bottom-up ML models are superior to top-down asset pricing models when the goal is to predict returns.

The remaining rows of Table VIII show the required pricing ability separately for required returns based on each of the six combinations of subjective risk and return expectations. Focusing on the first row that corresponds to Figure 7, we see that the \tilde{R}^2 of the traditional models ranges from 0.24 (CCAPM) to 0.63 (CAPM), while the recent models range from -5.25 to -3.20. The CAPM is the best model for five of six required return proxies, while CCAPM is the best for the remaining one.²⁵ FF3 is the second-best model for four of six proxies but is less stable than CCAPM.

The good performance of CCAPM is somewhat surprising considering the visual evidence from Figure 7. It reflects the low variance of the CCAPM predictions that counteracts its high bias. This mechanism also explains why the \tilde{R}^2 of CCAPM is so stable in Table VIII. As a practical matter, the high bias is undesirable for a good model of required returns, even if it comes at a low variance.

I develop a pairwise model comparison test to test whether the differences across models are statistically significant. Optimally, I would like to create a hypothesis test of whether one model is significantly better than another. However, the sampling distribution of (1.41) and (1.42) is complex, making it difficult to derive a null distribution. Instead, I base the model comparison on the confidence distribution of the difference estimated via a bootstrap procedure. A confidence distribution, $h(x)$, is closely related to confidence intervals. For example, for a central 90% confidence interval spanning x^{lower} to x^{high} , the lower endpoint satisfies $h(x^{lower}) = 0.05$ and the upper endpoint satisfies $h(x^{upper}) = 0.95$.

When comparing the realized pricing ability of two models, say the CAPM and FF3, I compute the bootstrap confidence distribution of $R_{CAPM}^2 - R_{FF3}^2$ and evaluate its value at zero. I refer to the resulting statistic as the bootstrap p -value. The bootstrap p -value, roughly, shows the proportion of time where R_{CAPM}^2 is lower than R_{FF3}^2 . I explain the

²⁵Perhaps surprisingly, CAPM has a negative \tilde{R}^2 when required returns are based on return expectation from Value Line and β_t^i . The reason is that the estimated subjective risk premium with these proxies is lower than implied by the CAPM. Said differently, Value Line believes that the relation between beta and expected returns is weaker than predicted by the CAPM.

Table VIII. Required and Realized Pricing Ability

Required ret. data	REQ	CAPM	CCAPM	FF3	FF5	HXZ	SY	DHS	ML
Panel A. Realized pricing ability (R^2)									
	-0.91	-0.92	-0.05	-0.36	0.26	0.42	0.45	0.32	0.78
Panel B. Required pricing ability (\tilde{R}^2)									
<i>VL</i> / $s_t^{VL,i}$	1.00	0.63	0.24	0.32	-3.20	-3.66	-5.25	-5.10	-3.59
<i>VL</i> / β_t^i	1.00	-0.33	0.37	-1.00	-9.92	-10.31	-16.23	-13.76	-9.82
<i>VL</i> / $SVIX_t^i$	1.00	0.66	0.25	-0.32	-5.63	-5.91	-9.03	-7.86	-5.42
<i>IBES</i> / $s_t^{VL,i}$	1.00	0.59	0.13	0.37	-1.15	-1.41	-1.93	-2.02	-1.42
<i>IBES</i> / β_t^i	1.00	0.79	0.22	0.25	-2.91	-3.17	-4.89	-4.29	-3.08
<i>IBES</i> / $SVIX_t^i$	1.00	0.79	0.14	0.19	-1.97	-2.14	-3.19	-2.88	-1.96

Note: The table shows the realized and required pricing ability of the nine asset pricing models from Table VII on the 357 characteristic sorted portfolios. Panel A shows the realized pricing ability computed using (1.41), and Panel B shows the required pricing ability computed using (1.42). The “Required ret. data” column refers to the subjective expected return (before the slash), and the subjective risk proxy (after the slash) used to compute required returns. The subjective expected return either comes from Value Line (VL) or I/B/E/S (IBES). The subjective risk is either the safety rank from Value Line ($s_t^{VL,i}$), the market beta from Value Line (β_t^i), or the SVIX measure from Martin and Wagner (2019). The model abbreviations refer to a required return model (REQ), classical capital asset pricing model (CAPM), the consumption CAPM (CCAPM), the Fama-French three-factor (FF3), the Fama-French five-factor model (FF5), the Hou-Xue-Zhang investment CAPM (HXZ), the Stambaugh-Yuan model with mispricing factors (SY), the Daniel-Hirshleifer-Sun model with behavioral factors (DHS), and a machine learning based model. The realized pricing ability of REQ is an average across the six required return proxies.

procedure in detail in Section 8.9.

Panel A in Table IX shows the results of the pairwise model comparisons for the realized pricing ability. A number is written in bold if the bootstrap p value is below 0.05, which I refer to as a significant difference. Table VIII showed that SY had the highest realized pricing ability, but the difference relative to FF5, HXZ, and DHS is not statistically significant. Within the four recent models, I only reject FF5 in favor of HXZ. In contrast, when comparing traditional and recent models, I almost always reject the traditional models in favor of the recent models. Overall, these results warrant the conclusion that recent empirical asset pricing models are significantly better at predicting returns than traditional models.

Table IX also reveals that the ML model is superior to all other models for predicting realized returns. Again, this evidence favors using ML rather than the recent asset pricing models for the “realized return” mandate.

Panel B in Table IX presents the model comparison for the required pricing ability using the safety rank and subjective expected return from Value Line. The CAPM is significantly better than all models except FF3. Even relative to FF3, the evidence favors the CAPM, with a bootstrap p -value of 0.22. Generally, the evidence favoring the CAPM is strong across the six required return proxies with a median bootstrap p -value of 0.02

Table IX. Recent models are significantly better for realized returns while traditional models are significantly better for required returns

Panel A: Test of R^2 difference for realized return

	REQ	CAPM	CCAPM	FF3	FF5	HXZ	SY	DHS	ML
REQ		0.79	0.00	0.13	0.00	0.00	0.00	0.00	0.00
CAPM	0.21		0.00	0.01	0.00	0.00	0.00	0.00	0.00
CCAPM	1.00	1.00		0.88	0.15	0.02	0.01	0.07	0.00
FF3	0.87	0.99	0.12		0.00	0.00	0.00	0.00	0.00
FF5	1.00	1.00	0.85	1.00		0.04	0.15	0.35	0.01
HXZ	1.00	1.00	0.98	1.00	0.96		0.53	0.86	0.03
SY	1.00	1.00	0.99	1.00	0.85	0.47		0.82	0.01
DHS	1.00	1.00	0.93	1.00	0.65	0.14	0.18		0.00
ML	1.00	1.00	1.00	1.00	0.99	0.97	0.99	1.00	

Panel B: Test of \tilde{R}^2 difference for required returns

	REQ	CAPM	CCAPM	FF3	FF5	HXZ	SY	DHS	ML
REQ		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
CAPM	0.00		0.96	0.78	1.00	1.00	1.00	1.00	1.00
CCAPM	0.00	0.04		0.38	1.00	1.00	1.00	1.00	1.00
FF3	0.00	0.22	0.62		1.00	1.00	1.00	1.00	1.00
FF5	0.00	0.00	0.00	0.00		0.68	0.95	0.97	0.47
HXZ	0.00	0.00	0.00	0.00	0.32		0.93	0.88	0.44
SY	0.00	0.00	0.00	0.00	0.05	0.07		0.48	0.29
DHS	0.00	0.00	0.00	0.00	0.03	0.12	0.52		0.23
ML	0.00	0.00	0.00	0.00	0.53	0.56	0.71	0.77	

Note: The table shows the results of the pairwise model comparisons of pricing ability. The number states the bootstrap p -value, explained in Section 8.9. Roughly, the number shows the proportion of bootstrap samples where the pricing ability of the row model was better than the column model. Numbers below 5% indicate statistically significant evidence in favor of the column model and are highlighted in bold. In panel A, the pricing metric is the R^2 from (1.41) and in of panel B the pricing metric is \tilde{R}^2 from (1.42). Required returns are based on $\tilde{E}_t^{VL}[r_{t,t+48}^i]$ and $s_t^{VL,i}$.

relative to CCAPM and 0.1 relative to FF3.

At a higher level, I reject all the recent models in favor of the traditional model. These results motivate the claim that traditional models are superior for explaining required returns.

7 Conclusion: A Risk-Return Tradeoff Concealed by Optimism Bias

I use subjective risk and return expectations to infer required returns via three different proxies for subjective risk and two for subjective expected returns. I compare these ex-ante required returns to ex-post realized returns in the cross-section of stocks.

My first finding is that the subjective risk premium is high while the realized risk premium is low. This stylized fact suggests that investors require substantial compensation for taking risks, but, in 30 years of data, the realized risk compensation has been disappointing.

I show theoretically that the realized risk premium is lower than the subjective risk premium when investors suffer from optimism bias. In particular, the expectational error induced by optimism bias increases in cash flow uncertainty, and empirically, riskier stocks have more uncertain cash flows. Using subjective cash flow forecasts from Value Line and I/B/E/S, I find strong support for three distinct predictions of optimism bias: (i) cash flow forecasts are too optimistic on average, (ii) over-optimism is more common for riskier stocks, and (iii) over-optimism is more common over longer forecast horizons. These results suggest that the high required return of risky stocks is offset by irrational cash flow optimism leading to a low realized return.

The weak link between realized and required returns has important implications for asset pricing. For equity factors, I test the “risk hypothesis” that a factor’s realized return is equal to its required return. In a sample of 119 factors, I reject the risk hypothesis for 71% to 79%. This high rejection rate reflects that most factors, even though they have a positive realized return, tend to buy subjective safe stocks while shorting risky ones.

For asset pricing models, I find that recent empirical models, such as FF5, are better than traditional models, such as the CAPM, at explaining realized returns. Conversely, the recent models perform poorly in explaining required returns, whereas the traditional models perform well. The CAPM, in particular, emerges as the leading model of required returns.

My results suggest that realized returns contain limited information about investors’ required returns because of the relation between risk and irrational cash flow optimism.

8 Appendix

8.1 Proofs

I start by showing that the objective expected return under the setup in Section 1 is,

$$\mathbb{E}_t[r_{t+1}^i] = -\gamma_t^0 + (\lambda_t - \gamma_t^1)s_t^i - u_t^i, \quad (1.43)$$

which follows from the definition of objective mispricing as $b_t^i = \lambda s^i - \mathbb{E}_t[r_{t+1}^i]$ and the decomposition of b_t^i from (1.5). Next, γ_t^0 and γ_t^1 are the coefficients from a cross-sectional regression of objective mispricing on subjective risk leading, so their expected values are,

$$\gamma_t^0 = \bar{b}_t - \gamma_1 \bar{s}_t, \quad (1.44)$$

$$\gamma_t^1 = \frac{\overline{\text{Cov}}(s_t^i, b_t^i)}{\overline{\text{Var}}(s_t^i)}, \quad (1.45)$$

where \bar{b}_t is the average objective bias at time t , \bar{s}_t is the average subjective risk, and $\overline{\text{Var}}$ and $\overline{\text{Cov}}$ refers to cross-sectional variance and covariance, respectively.

Proof of Proposition 1. The λ_t estimate from (1.6) is

$$\begin{aligned} \hat{\lambda}_t &= \frac{\overline{\text{Cov}}(\tilde{\mathbb{E}}_t[r_{t+1}^i], s_t^i)}{\overline{\text{Var}}(s^i)} \\ &= \frac{\overline{\text{Cov}}(\lambda_t s_t^i - \tilde{b}_t^i, s_t^i)}{\overline{\text{Var}}(s^i)} \\ &= \lambda_t - \frac{\overline{\text{Cov}}(\tilde{b}_t^i, s_t^i)}{\overline{\text{Var}}(s^i)}, \end{aligned}$$

where the second line expands $\tilde{\mathbb{E}}_t[r_{t+1}^i]$ according to (1.4).

Next, the λ_t^{rea} coefficient from (1.7) is

$$\begin{aligned} \lambda_t^{rea} &= \frac{\overline{\text{Cov}}(\mathbb{E}_t[r_{t+1}^i], s_t^i)}{\overline{\text{Var}}(s^i)} \\ &= \frac{\overline{\text{Cov}}(\lambda_t s_t^i - b_t^i, s_t^i)}{\overline{\text{Var}}(s^i)} \\ &= \lambda_t - \frac{\overline{\text{Cov}}(b_t^i, s_t^i)}{\overline{\text{Var}}(s^i)} \\ &= \lambda_t - \gamma_t^1, \end{aligned}$$

where the second line follows from the definition of objective expected returns in (1.43) and the last line follows from the definition of γ_t^1 as the coefficient from a cross-sectional

regression of objective bias on subjective risk (see (1.5)). ■

Proof of Proposition 2. Since the expected return of a portfolio is linear in the portfolio assets, we can use (1.43) to get

$$\begin{aligned}
\mathbb{E}_t[r_{t+1}^L - r_{t+1}^S] &= \sum_{i \in L} \pi_t^i (-\gamma_t^0 + (\lambda_t - \gamma_t^1) s_t^i - u_t^i) - \sum_{i \in S} \pi_t^i (-\gamma_t^0 + (\lambda_t - \gamma_t^1) s_t^i - u_t^i), \\
&= \sum_{i \in L} \pi_t^i (\lambda_t s_t^i - \gamma_t^1 s_t^i - u_t^i) - \sum_{i \in S} \pi_t^i (\lambda_t s_t^i - \gamma_t^1 s_t^i - u_t^i) \\
&= \lambda_t \left(\sum_{i \in L} \pi_t^i s_t^i - \sum_{i \in S} \pi_t^i s_t^i \right) - \gamma_t^1 \left(\sum_{i \in L} \pi_t^i s_t^i - \sum_{i \in S} \pi_t^i s_t^i \right) - \left(\sum_{i \in L} \pi_t^i u_t^i - \sum_{i \in S} \pi_t^i u_t^i \right) \\
&= \lambda_t (s_t^L - s_t^S) - \gamma_t^1 (s_t^L - s_t^S) - (u_t^L - u_t^S)
\end{aligned}$$

The second line removes all constants because the weight in both the long and short portfolio sum to 1. The third line collects terms that depend on the same parameters, and the fourth line defines s_t^j and u_t^j . ■

Proof of Proposition 3. I start by deriving the values of κ_t^0 and κ_t^1 . The two parameters come from a cross-sectional regression of objective expected returns on a constant and subjective risk,

$$\mathbb{E}_t[r_{t+1}^i] = \kappa_t^0 + \kappa_t^1 s_t^i + \epsilon_t^i$$

Starting with κ_t^1 :

$$\begin{aligned}
\kappa_t^1 &= \frac{\overline{\text{Cov}}(\mathbb{E}_t[r_{t+1}^i], s_t^i)}{\overline{\text{Var}}(s_t^i)} \\
&= \frac{\overline{\text{Cov}}(-\gamma_t^0 + (\lambda_t - \gamma_t^1) s_t^i + u_t^i, s_t^i)}{\overline{\text{Var}}(s_t^i)} \\
&= (\lambda_t - \gamma_t^1) \frac{\overline{\text{Cov}}(s_t^i, s_t^i)}{\overline{\text{Var}}(s_t^i)} \\
&= \lambda_t - \gamma_t^1,
\end{aligned} \tag{1.46}$$

where the second line expands $\mathbb{E}_t[r_{t+1}^i]$ according to (1.43), and the fourth line follows because u_t^i and s_t^i are cross-sectionally uncorrelated and all the remaining terms are constant. Next, κ_t^0 follows as

$$\kappa_t^0 = \mathbb{E}_t[r_{t+1}^i] - \gamma_t^1 \bar{s}_t - \bar{b}_t \tag{1.47}$$

Objective pricing error

I start by defining the objective pricing error for a single asset as

$$\alpha_m^{(i)} = \mathbb{E}_t[r_{t+1}^i] - \mathbb{E}_t^m[r_{t+1}^i]. \tag{1.48}$$

1. *Required return model, (1.13)*

Subtracting (1.13) from the objective expected return, (1.43), we see that the required return model cancels $\lambda_t s_t^i$ such that

$$\alpha_{\text{req}}^{(i)} = -\gamma_t^0 - \gamma_t^1 s_t^i - u_t^i.$$

The objective pricing error of this model is therefore

$$\begin{aligned} \alpha_{\text{req}}^2 &= \mathbb{E}[(-\gamma_t^0 - \gamma_t^1 s_t^i - u_t^i)^2] \\ &= \mathbb{E}[(-\bar{b}_t - \gamma_t^1 \bar{s}_t) - \gamma_t^1 s_t^i - u_t^i]^2 \\ &= \mathbb{E}[(-\bar{b}_t - \gamma_t^1 (s_t^i - \bar{s}_t) - u_t^i)^2] \\ &= \mathbb{E}_t[-\bar{b}_t - \gamma_t^1 (s_t^i - \bar{s}_t) - u_t^i]^2 + \overline{\text{Var}}(-\bar{b}_t - \gamma_t^1 (s_t^i - \bar{s}_t) - u_t^i) \\ &= \bar{b}_t^2 + (\gamma_t^1)^2 \overline{\text{Var}}(s_t^i) + \overline{\text{Var}}(u_t^i), \end{aligned}$$

where the second line expands γ_t^0 according to (1.44), so \bar{b}_t is the average objective mispricing at time t . The fifth line follows because s_t^i and u_t^i is cross-sectionally uncorrelated.

2. *Empirical free model, (1.11)* The pricing error at the asset level is

$$\begin{aligned} \alpha_{\text{free}}^{(i)} &= \mathbb{E}_t[r_{t+1}^i] - (\kappa_t^0 + \kappa_t^1 s_t^i) \\ &= -\gamma_t^0 + (\lambda_t - \gamma_t^1) s_t^i - u_t^i - (-\gamma_t^0 + (\lambda_t - \gamma_t^1) s_t^i) \\ &= -u_t^i, \end{aligned}$$

where the second line follows from the definition of the objective expected return from (1.43) and I expand κ_t^1 and κ_t^0 using (1.46) and (1.47), respectively. The model's overall objective pricing error is

$$\alpha_{\text{free}}^2 = \mathbb{E}[-u_t^i]^2 + \overline{\text{Var}}(u_t^i) = \overline{\text{Var}}(u_t^i).$$

this error is smaller than or equal to α_{req}^2 because $\alpha_{\text{req}}^2 - \alpha_{\text{free}}^2 = \bar{b}_t^2 + (\gamma_t^1)^2 \overline{\text{Var}}(s_t^i)$ is non-negative.

3. *Empirical multifactor model, (1.12)*

The pricing error at the asset level is

$$\alpha_{\text{multifactor}}^{(i)} = -u_t^i - \kappa_2 c_t^i,$$

which is the same as the empirical free model, except that the multifactor model

also tackles non-risk mispricing (u_t^i). The model's overall objective pricing error is

$$\begin{aligned}
\alpha_{\text{multifactor}}^2 &= \text{E}_t[(-u_t^i - \kappa_2 c_t^i)^2] \\
&= \overline{\text{Var}}_t(u_t^i) + \kappa_2^2 \overline{\text{Var}}_t(c_t^i) + 2\kappa_2 \overline{\text{Cov}}_t(u_t^i, c_t^i) \\
&= \overline{\text{Var}}_t(u_t^i) + \left(\frac{-\overline{\text{Cov}}_t(u_t^i, c_t^i)}{\overline{\text{Var}}_t(c_t^i)} \right)^2 \overline{\text{Var}}_t(c_t^i) - 2 \frac{\overline{\text{Cov}}_t(u_t^i, c_t^i)}{\overline{\text{Var}}_t(c_t^i)} \overline{\text{Cov}}_t(u_t^i, c_t^i) \\
&= \overline{\text{Var}}_t(u_t^i) - \frac{\overline{\text{Cov}}_t(u_t^i, c_t^i)^2}{\overline{\text{Var}}_t(c_t^i)}, \\
&= \overline{\text{Var}}_t(u_t^i) [1 - \overline{\text{Cor}}_t(u_t^i, c_t^i)^2],
\end{aligned}$$

where the second line uses the assumption that both u_t^i and c_t^i are mean zero, the third line uses the definition of $\kappa_2 = \overline{\text{Cov}}(\text{E}_t[r_{t+1}^i], c_t^i) / \overline{\text{Var}}(c_t^i) = -\overline{\text{Cov}}(u_t^i, c_t^i) / \overline{\text{Var}}(c_t^i)$, and $\overline{\text{Cor}}$ is the cross-sectional correlation. This pricing error is less than or equal to α_{free}^2 because the squared correlation is non-negative.

4. Objective model, (1.14)

Since $\text{E}_t^{\text{obj}}[r_t^i] = \text{E}_t[r_{t+1}^i], \forall i$, the objective model has a zero objective pricing error by definition,

$$\alpha_{\text{obj}}^2 = 0.$$

This error is the lowest possible because (1.15) is bounded below by zero (since the square of a real number is non-negative).

Based on the derivations above, the first result in Proposition 3 follows

$$\underbrace{\alpha_{\text{req}}^2}_{\overline{b}_t^2 + (\gamma_t^1)^2 \overline{\text{Var}}(s_t^i) + \overline{\text{Var}}(u_t^i)} \geq \underbrace{\alpha_{\text{free}}^2}_{\overline{\text{Var}}(u_t^i)} \geq \underbrace{\alpha_{\text{multifactor}}^2}_{\overline{\text{Var}}_t(u_t^i) [1 - \overline{\text{Cor}}_t(u_t^i, c_t^i)^2]} \geq \alpha_{\text{obj}}^2 = 0.$$

Subjective pricing error

Again, I start by defining the subjective pricing error for a single asset as

$$\tilde{\alpha}_m^{(i)} = \lambda_t s_t^i - \text{E}_t^m[r_{t+1}^i]. \quad (1.49)$$

1. Required return model, (1.13)

Since $\text{E}_t^{\text{req}}[r_{t+1}^i] = \lambda_t s_t^i, \forall i$, the required return model has a zero subjective pricing error by definition,

$$\tilde{\alpha}_{\text{req}}^2 = 0.$$

This error is the lowest possible because (1.16) is bounded below by zero (since the square of a real number is non-negative).

2. *Empirical free model, (1.11)*

The subjective pricing error at the asset level is

$$\begin{aligned}\tilde{\alpha}_{\text{free}}^{(i)} &= \lambda_t s_t^i - (\kappa_t^0 + \kappa_t^1 s_t^i) \\ &= \lambda_t s_t^i - (\gamma_t^1 \bar{s}_t - \bar{b}_t + (\lambda_t - \gamma_t^1) s_t^i) \\ &= \bar{b}_t + \gamma_t^1 (s_t^i - \bar{s}_t),\end{aligned}$$

where the second line expands κ_t^1 and κ_t^0 according to (1.46) and (1.47), respectively. The overall subjective pricing error of this model is therefore

$$\begin{aligned}\tilde{\alpha}_{\text{free}}^2 &= \text{E}[(\bar{b}_t + \gamma_t^1 (s_t^i - \bar{s}_t))^2] \\ &= \bar{b}_t^2 + (\gamma_t^1)^2 \overline{\text{Var}}(s_t^i),\end{aligned}$$

which is larger than or equal to $\tilde{\alpha}_{\text{req}}^2$ because $\bar{b}_t^2 + (\gamma_t^1)^2 \overline{\text{Var}}(s_t^i)$ is non-negative.

3. *Empirical multifactor model, (1.12)*

The subjective pricing error at the asset level is

$$\begin{aligned}\tilde{\alpha}_{\text{multifactor}}^{(i)} &= \lambda_t s_t^i - (\kappa_t^0 + \kappa_t^1 s_t^i + \kappa_t^2 c_t^i) \\ &= \bar{b}_t + \gamma_t^1 (s_t^i - \bar{s}_t) - \kappa_t^2 c_t^i\end{aligned}$$

which is the same as the empirical free model, except for the additional term, $\kappa_t^2 c_t^i$. The overall subjective pricing error of the empirical multifactor is therefore

$$\begin{aligned}\tilde{\alpha}_{\text{multifactor}}^2 &= \text{E}[(\bar{b}_t + \gamma_t^1 (s_t^i - \bar{s}_t) - \kappa_t^2 c_t^i)^2] \\ &= \bar{b}_t^2 + (\gamma_t^1)^2 \overline{\text{Var}}(s_t^i) + \kappa_t^2 \overline{\text{Var}}(c_t^i) \\ &= \bar{b}_t^2 + (\gamma_t^1)^2 \overline{\text{Var}}(s_t^i) + \overline{\text{Cor}}(u^i, c_t^i)^2 \overline{\text{Var}}(u_t^i),\end{aligned}$$

which is larger than or equal to $\tilde{\alpha}_{\text{free}}^2$ because $\overline{\text{Cor}}(u^i, c_t^i)^2$ and $\overline{\text{Var}}(u_t^i)$ are both non-negative.

4. *Objective model, (1.14)*

The subjective pricing error at the asset level is

$$\begin{aligned}\tilde{\alpha}_{\text{obj}}^{(i)} &= \lambda_t s_t^i - (\kappa_t^0 + \kappa_t^1 s_t^i + u_t^i) \\ &= \bar{b}_t + \gamma_t^1 (s_t^i - \bar{s}_t) - u_t^i.\end{aligned}$$

The subjective pricing error of this model is therefore

$$\begin{aligned}\tilde{\alpha}_{\text{obj}}^2 &= \mathbb{E}[(\bar{b}_t + \gamma_t^1(s_t^i - \bar{s}_t) - u_t^i)^2] \\ &= \bar{b}_t^2 + (\gamma_t^1)^2 \overline{\text{Var}}(s_t^i) + \overline{\text{Var}}(u_t^i),\end{aligned}$$

which is larger than or equal to $\tilde{\alpha}_{\text{multifactor}}^2$ because $\overline{\text{Cor}}(u^i, c_t^i)^2 \leq 1$.

Based on the derivations above, the second result in Proposition 3 follows

$$\underbrace{\tilde{\alpha}_{\text{obj}}^2}_{\bar{b}_t^2 + (\gamma_t^1)^2 \overline{\text{Var}}(s_t^i) + \overline{\text{Var}}(u_t^i)} \geq \underbrace{\tilde{\alpha}_{\text{multifactor}}^2}_{\bar{b}_t^2 + (\gamma_t^1)^2 \overline{\text{Var}}(s_t^i) + \overline{\text{Cor}}(u^i, c_t^i)^2 \overline{\text{Var}}(u_t^i)} \geq \underbrace{\tilde{\alpha}_{\text{free}}^2}_{\bar{b}_t^2 + (\gamma_t^1)^2 \overline{\text{Var}}(s_t^i)} \geq \tilde{\alpha}_{\text{req}}^2 = 0.$$

■

Proof of Proposition 4. I start by deriving the rational posterior expectation from (1.17). The three random variables of interest are the expected cash flow $\theta_i \sim N(\mu_0, \tau_0^2 \omega_i^2)$ and the two public signals $v_i^k | \theta_i \sim N(\theta_i, \tau_1^2 \omega_i^2)$ with $k \in \{1, 2\}$. These three variables have a multivariate normal distribution if every linear combination $Y = a\theta_i + bs_1^i + cs_2^i$ of its components is normally distributed. To see that this is the case, I write each signal as the sum of theta and a random variable u_i^k : $Y = a\theta_i + b(\theta_i + u_i^1) + c(\theta_i + u_i^2) = \theta_i(a + b + c) + bu_i^1 + cu_i^2$. Hence, Y is a linear combination of three independent normally distributed variables, which is itself normally distributed. The prior distribution of the expected cash flows and the two public signals are, therefore, multivariate normal:

$$\begin{pmatrix} \theta_i \\ v_i^1 \\ v_i^2 \end{pmatrix} = N \left(\begin{bmatrix} \mu_0 \\ \mu_0 \\ \mu_0 \end{bmatrix}, \omega_i^2 \begin{bmatrix} \tau_0^2 & \tau_0^2 & \tau_0^2 \\ \tau_0^2 & \tau_0^2 + \tau_1^2 & \tau_0^2 \\ \tau_0^2 & \tau^2 & \tau_0^2 + \tau_1^2 \end{bmatrix} \right) \quad (1.50)$$

From the properties of a multivariate normal distribution, the conditional distribution is:

$$\begin{aligned}E[\theta_i | v_i^1, v_i^2] &= \mu_0 + \frac{\omega_i^2}{\omega_i^2} \begin{bmatrix} \tau_0^2 & \tau_0^2 \end{bmatrix} \begin{bmatrix} \tau_0^2 + \tau_1^2 & \tau^2 \\ \tau^2 & \tau_0^2 + \tau_1^2 \end{bmatrix}^{-1} \begin{bmatrix} v_i^1 - \mu_0 \\ v_i^2 - \mu_0 \end{bmatrix} \\ &= \mu_0 + \frac{1}{(\tau_0^2 + \tau_1^2)(\tau_0^2 + \tau_1^2) - \tau_0^4} \begin{bmatrix} \tau_0^2 & \tau_0^2 \end{bmatrix} \begin{bmatrix} \tau_0^2 + \tau_1^2 & -\tau_0^2 \\ -\tau_0^2 & \tau_0^2 + \tau_1^2 \end{bmatrix} \begin{bmatrix} v_i^1 - \mu_0 \\ v_i^2 - \mu_0 \end{bmatrix} \\ &= \mu_0 + \frac{1}{(\tau_0^2 + \tau_1^2)(\tau_0^2 + \tau_1^2) - \tau_0^4} \begin{bmatrix} \tau_0^2(\tau_0^2 + \tau_1^2 - \tau_0^2) & \tau_0^2(\tau_0^2 + \tau_1^2 - \tau_0^2) \end{bmatrix} \begin{bmatrix} v_i^1 - \mu_0 \\ v_i^2 - \mu_0 \end{bmatrix} \\ &= \mu_0 + \frac{1}{2} \frac{\tau_0^2}{\tau_0^2 + \tau_1^2/2} (v_i^1 - \mu_0) + \frac{1}{2} \frac{\tau_0^2}{\tau_0^2 + \tau_1^2/2} (v_i^2 - \mu_0) \\ &= \mu_0 + \delta \left(\frac{1}{2} v_i^{\text{max}} + \frac{1}{2} v_i^{\text{min}} - \mu_0 \right),\end{aligned}$$

where $\delta = \frac{\tau_0^2}{\tau_0^2 + \tau_1^2/2}$ is a shrinkage parameter which is the same for all stocks, $v_i^{max} := \max(v_i^1, v_i^2)$ is the highest signal, and $v_i^{min} := \min(v_i^1, v_i^2)$ is the lowest signal. Hence, the rational posterior is an equal-weighted average of the two signals.

Next, I want to show the expected bias when the investors actual posterior expectation puts κ^j weight on v_i^{max} and $1 - \kappa_j$ on v_i^{min} :

$$E^j[\theta_i | v_i^{max}, v_i^{min}, \kappa^j] = \mu_0 + \delta(\kappa_j v_i^{max} + (1 - \kappa_j) v_i^{min} - \mu_0)$$

Define the bias in investor inference as $b_j^i = E^j[\theta_i | v_i^{max}, v_i^{min}, \kappa^j] - E[\theta_i | v_i^{max}, v_i^{min}]$. This bias depends on the two random signals, but I want to characterize its expected value.

To do so, I use the results from Nadarajah and Kotz (2008) about the expected value of the maximum and minimum of two normal random variables. If X_1 and X_2 are jointly normal with expected value μ_X and μ_Y , variances σ_X^2 and σ_Y^2 , and covariance $cov(X, Y)$, then their expected values are:

$$E[\max(X_1, X_2)] = \mu_X \Phi\left(\frac{\mu_X - \mu_Y}{\xi}\right) + \mu_Y \Phi\left(\frac{\mu_Y - \mu_X}{\xi}\right) + \xi \phi\left(\frac{\mu_X - \mu_Y}{\xi}\right),$$

$$E[\min(X_1, X_2)] = \mu_X \Phi\left(\frac{\mu_X - \mu_Y}{\xi}\right) + \mu_Y \Phi\left(\frac{\mu_Y - \mu_X}{\xi}\right) - \xi \phi\left(\frac{\mu_X - \mu_Y}{\xi}\right),$$

where $\Phi(x)$ is the cumulative distribution function of a standard normal variable, $\phi(x)$ is the density function of a standard normal variable, and $\xi = \sqrt{\sigma_X^2 + \sigma_Y^2 - 2Cov(X, Y)}$. If the two variables have the same expected mean μ and variance, σ^2 , then this expression reduces to:

$$E[\max(X_1, X_2)] = \mu + \xi \phi(0) \tag{1.51}$$

$$E[\min(X_1, X_2)] = \mu - \xi \phi(0) \tag{1.52}$$

Using this result, I can derive the expected bias from Proposition 4:

$$\begin{aligned} E[b_j^i] &= E[\mu_0 + \delta(\kappa_j v_i^{max} + (1 - \kappa_j) v_i^{min} - \mu_0) - (\mu_0 + \delta(0.5 v_i^{max} + 0.5 v_i^{min} - \mu_0))] \\ &= \delta(E[v_i^{max}] - E[v_i^{min}])(\kappa_j - 0.5) \\ &= \delta 2\phi(0) \xi_i (\kappa_j - 0.5) \\ &= \delta 2\phi(0) \sqrt{2\omega_i^2(\tau_0^2 + \tau_1^2) - 2\omega_i^2\tau_0^2} (\kappa_j - 0.5) \\ &= \delta 2^{1.5} \tau_1 \phi(0) (\kappa_j - 0.5) \omega_i \\ &= c(\kappa_j - 0.5) \omega_i, \end{aligned}$$

where $\Phi(x)$ is the cumulative distribution function of a standard normal variable, $\phi(x)$ is the density function of a standard normal variable, and $c = \delta 2^{1.5} \tau_1 \phi(0)$ is a constant. On the third line, I evaluate $E[v_i^{max}]$ and $E[v_i^{min}]$ using (1.51) and (1.52), respectively; on the fourth line, I use the variance and covariance from (1.50); and on the final line I define c . Furthermore, c is positive because I assume that τ_0 and τ_1 are positive (meaning that δ is positive), and the remaining components are positive constants. ■

8.2 Subjective Risk and Cash Flow Uncertainty

A critical assumption for Proposition 4, is that cash flow uncertainty is higher for subjectively riskier stocks. The Value Line data provides excellent data for testing this assumption. Specifically, each reports contains an “earnings predictability” score ranging between 0 (low predictability) and 100 (high predictability).²⁶ To test whether subjective risk correlates with economic uncertainty, I sort stocks into 10 portfolio based on subjective risk with monthly re-balancing. The portfolios are identical to the ones in figure 1. For each portfolio-month, I compute the average earnings predictability scores. I then average these scores over time. Computing the standard error of this estimate is complicated by the fact that earnings predictability is persistent.²⁷ To circumvent this issue, I use an adjustment from (Cochrane, 2005, p.223) which is valid if earnings predictability at the portfolio level has an AR(1) structure,

$$SE(\bar{x}_k) = \sqrt{\frac{\text{Var}(x_k)}{T} \times \frac{1 + \rho_k}{1 - \rho_k}}, \quad (1.53)$$

where x_k is the earning predictability in portfolio k , T is the number of time-periods, and ρ_k is the monthly autocorrelation of x_k . Figure A1 shows that riskier stocks have higher cash flow uncertainty (lower earnings predictability) across all three risk proxies. Focusing on the safety rank from Value Line (s^{VL}), the safest stocks have an average earnings predictability of 83 while the corresponding number is 23 for the riskiest stock. Both numbers are estimated very precisely. Furthermore the relationship is monotonic across risk groups. Across stocks and time-periods, the correlation with earnings predictability is -0.61, -0.36, and -0.46 for the safety rank, beta estimate, and SVIX, respectively. These results support the assumption that cash flow uncertainty is higher for subjectively riskier assets.

²⁶The earnings predictability score is in the lower right corner of the report on Apple in figure A2.

²⁷The average 1 month autocorrelation for the 10 portfolios is 0.91.

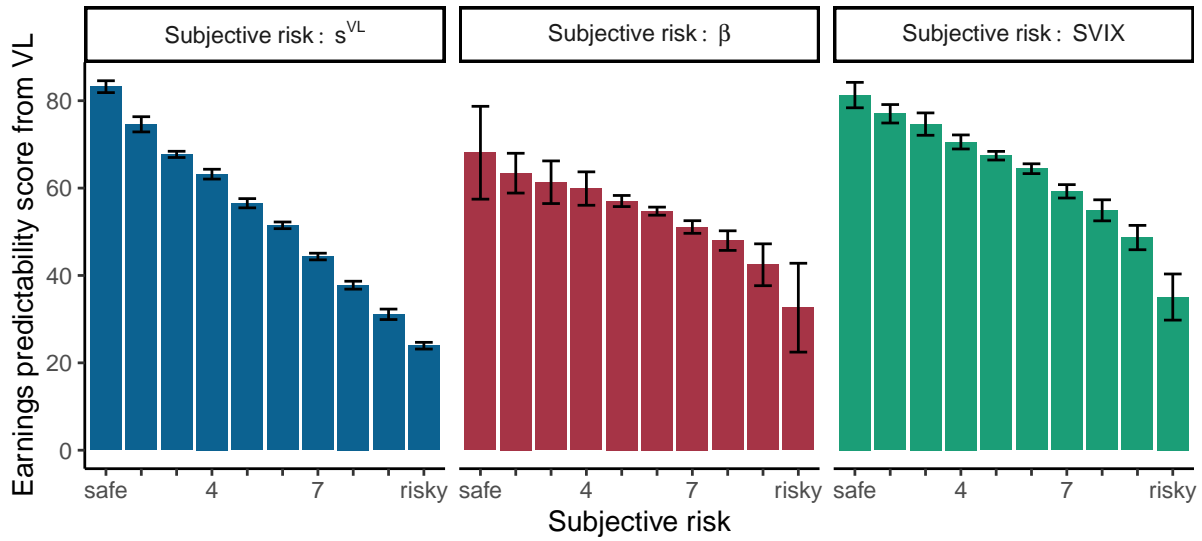


Figure A1. Subjective Risk and Cash Flow Uncertainty

Note: The figure shows the average earnings predictability score of stocks sorted into portfolio based on their subjective risk (1=safe, 10=risky). The earnings predictability score is from Value Line and it ranges from 0 (low predictability) to 100 (high predictability) and the score. The error bars show 95% confidence intervals computed as $\pm 1.96 \times SE(\bar{x}_k)$, where $SE(\bar{x}_k)$ is defined in (1.53). The subjective risk proxies are either the safety rank from Value Line, the market beta forecast from Value Line, or the SVIX measure from Martin and Wagner (2019).

8.3 Example of Value Line Investment Report

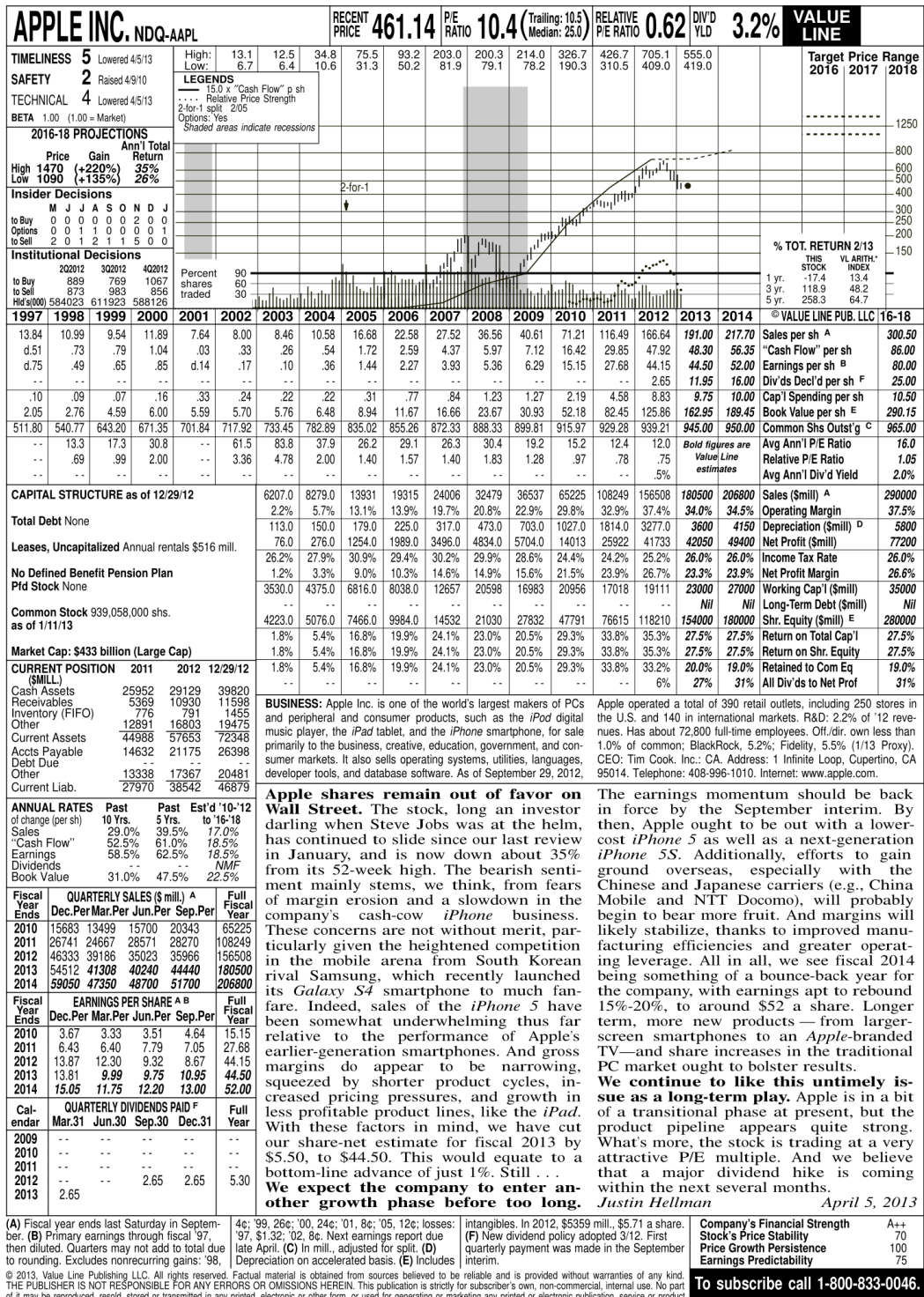


Figure A2. Value Line Investment Report

Note: The figure shows an example of a report from the Value Line Investment Survey for Apple. The example also displays most of the information from the data set used in this paper. For example, the key subjective risk measure is the safety rank in the top left corner. I transform the safety rank into a continuous measure by taking an average of its subcomponent, financial strength, and price stability shown in the bottom left corner. The primary input for the subjective return expectation is the price range projections visible just below the safety rank box.

8.4 Subjective Expected Return Calculation

Value Line

Value Line provides a price target four years ahead, $\tilde{E}_t[p_{t+4}^i]$, dividend expectations over the next calendar year, $\tilde{E}_t[d_{t+1}^i]$, and in four years, $\tilde{E}_t[d_{t+4}^i]$. I use this information and the current price, p_{t+1}^i , to calculate Value Line's implied return expectation over the next four years. For notational convenience, I suppress the stock identifier i . Furthermore, I denote the total expected return over the next year as $\tilde{E}_t[r_{t+1}^*] = r_{t+1}^i + r_{t+1}^f = \frac{\tilde{E}_t[p_{t+1}^i + d_{t+1}^i]}{p_t}$. The objective is to find the value of $\tilde{E}_t[r_{t+1}^*]$, consistent with Value Line's forecast and the stock's current price. The specific procedure depends on whether a firm has non-zero dividend expectations.

For firms where the two dividend expectations are both zero, the subjective return expectation is

$$\tilde{E}_t[r_{t+1}^*] = \left(\frac{\tilde{E}_t[p_{t+4}]}{p_t} \right)^{1/4} - 1.$$

For firms where the expected dividend in one year is zero, but the expectation in four years is positive, I assume that the expected dividend grows linearly over time, $\tilde{E}_t[d_{t+1+k}] = \frac{k}{3}\tilde{E}_t[d_{t+4}]$. The subjective returns expectation is the value of $\tilde{E}_t[r_{t+1}^*]$ that solves

$$(1 + \tilde{E}_t[r_{t+1}^*])^4 = \frac{\tilde{E}_t[p_{t+4}]}{p_t} + \frac{\tilde{E}_t[d_{t+2}](1 + \tilde{E}_t[r_{t+1}^*])^2 + \tilde{E}_t[d_{t+3}](1 + \tilde{E}_t[r_{t+1}^*]) + \tilde{E}_t[d_{t+3}]}{p_t}$$

For firms where both dividend expectations are non-zero, the subjective return expectation is the value of $\tilde{E}_t[r_{t+1}^*]$ that solves

$$(1 + \tilde{E}_t[r_{t+1}^*])^4 = \frac{\tilde{E}_t[p_{t+4}]}{p_t} + \frac{\tilde{E}_t[d_{t+1}]}{p_t} \frac{(1 + \tilde{E}_t[r_{t+1}^*])^4 - (1 + \tilde{E}_t[g_{t+4}])^4}{\tilde{E}_t[r_{t+1}^*] - \tilde{E}_t[g_{t+4}]}$$

where $\tilde{E}_t[g_{t+4}]$ is the expected dividend growth from year $t+1$ to $t+4$ computed as

$$\tilde{E}_t[g_{t+4}] = \left(\frac{\tilde{E}_t[d_{t+4}]}{\tilde{E}_t[d_{t+1}]} \right)^{1/3} - 1.$$

The four year expected excess return then follow as,

$$\tilde{E}_t[r_{t,t+4}^i] = (1 + \tilde{E}_t[r_{t+1}^*])^4 - 1 - r_{t,t+4}^f, \quad (1.54)$$

where $r_{t,t+4}^f$ is the risk-free rate from year t to $t+4$.

Comment

Value Line provides their expected annualized total return in their reports. In figure A2 it

is visible in the top left corner. However, in the data I received, there appears to be some issue with this data item before 2000. Specifically, when I look at old reports, the value for the expected return does not match those I have in my data. In contrast, the price target and dividend expectations match. After 2000, my implied return expectations match their data item almost perfectly. To ensure consistency, I compute the implied expected return target throughout the sample.

I/B/E/S

The subjective expected return of a stocks from I/B/E/S is,

$$\tilde{E}_t^{\text{IBES}}[r_{t+1}^i] = \frac{\tilde{E}_t^{\text{IBES}}[p_{t+1}^i] + \tilde{E}_t^{\text{IBES}}[d_{t+1}^i]}{p_t} - (1 + r_{t+1}^f), \quad (1.55)$$

where $\tilde{E}_t^{\text{IBES}}[p_{t+1}^i]$ and is the median consensus one-year price target from I/B/E/S, $\tilde{E}_t^{\text{IBES}}[d_{t+1}^i]$ is the median consensus dividend forecast over the next fiscal year from I/B/E/S, p_t^i is the stock's price at the day of the forecast, and r_{t+1}^f is the one-year risk-free rate. If the dividend forecast from I/B/E/S is unavailable, I use the one-year ahead dividend forecast from Value Line instead. If the Value Line forecast is also unavailable, I assume that the expected dividend is zero. Dividend expectations from I/B/E/S are only available for a broad cross-section of firms from 2002/05/16, so I use Value Line's dividend expectations before this date. The two series are highly similar as the Spearman correlation between the implied dividend yield from I/B/E/S and the dividend yield from Value Line is 0.96.

8.5 The Subjective Risk Premium is Countercyclical

To recover required returns from subjective risk and return expectations, there needs to be a zero correlation between subjective risk and subjective mispricing (Proposition 1). Throughout the paper, I use a constant subjective risk premium because I argue that the zero correlation assumptions is more likely to holds across on average across time-periods.

However, even if the zero correlation assumption is unlikely to hold exactly each month, the monthly parameter estimates can still be informative about the time-variation in required returns. Therefore, I use the time-varying coefficients from (1.21) based on the safety rank and expected return from Value Line, to compute the implied required return on the market portfolio.

The top-left panel in Figure A3 shows that the required market return is highly countercyclical, being high in bad times such as during the dotcom bubble in 2002-2003, the financial crisis in 2008, and the COVID crisis in 2020 and low in good times such as 2013-2019. The variation in the required return is large economically, from less than 5% to

more than 25%. These results suggest that the subjective risk premium is countercyclical.

Figure A3 also presents three alternatives. The top-right corner shows the risk aversion index from Bekaert et al. (2022).²⁸ This series represents an estimate of the relative risk aversion for an investor with habit preferences, and I take it as a proxy for “rational risk aversion.” The required market return from Value Line and the risk aversion index has a correlation of 0.69, suggesting that the subjective risk premium closely follows rational risk aversion proxies.

Next, Greenwood and Shleifer (2014) shows that the expected market return of many agents are pro-cyclical, so I also compare the required market return to the return expectations of different market participants. To proxy for the return expectations of professional investors, I use the Livingston survey of professional forecasters.²⁹ To proxy for the return expectations of retail investors, I use data from Nagel and Xu (2022a) based on various surveys of individuals.³⁰

The lower-left panel of Figure A3 shows that the return expectations of professional investors are very similar to the required market returns from Value Line, with a correlation of 0.70. In contrast, the return expectations of individual investors are very different from the required market return, with a correlation of -0.11. This finding suggests that the required return based on Value Line data is more representative of professional investors and less of retail investors.

8.6 Subjective Risk and Long-Term Growth Expectations

In Figure A4, I sort stocks into ten portfolios based on subjective risk, compute the average LTG forecast for each portfolio, and average this number over time. The left panel shows that I/B/E/S expects the safest decile of stocks to have an LTG below 10%, while the LTG of the riskiest stocks is around 18%. The right panel shows that this difference is even greater when using LTG forecasts from Value Line.

8.7 Equity Factor Information

Table AI. Factor Information

²⁸I extract the risk aversion directly from <https://www.nancyxu.net/risk-aversion-index>. I thank the authors for making the data available.

²⁹This data is maintained by the Philadelphia FED and is available at [<https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/livingston-survey>]. I use the ratio of the mean 12-month S&P500 forecast to the zero-month level minus one plus the 1-year risk-free rate. Dahlquist and Ibert (2021) show that the Livingston forecasts are countercyclical while Nagel and Xu (2022b) argue that the countercyclicality is smaller than implied by predictive regressions.

³⁰The return expectations of individual investors is available at <https://voices.uchicago.edu/stefannagel/code-and-data/>. I thank the authors for making the data available.

	Factor	Long	Theme
1	age	Low	Low Leverage
2	ami_126d	High	Size
3	at_gr1	Low	Investment
4	be_gr1a	Low	Investment
5	be_me	High	Value
6	beta_60m	Low	Low Risk
7	betabab_1260d	Low	Low Risk
8	betadown_252d	Low	Low Risk
9	bev_mev	High	Value
10	bidaskhl_21d	High	Low Leverage
11	capex_abn	Low	Debt Issuance
12	capx_gr2	Low	Investment
13	capx_gr3	Low	Investment
14	chcsho_12m	Low	Value
15	coa_gr1a	Low	Investment
16	col_gr1a	Low	Investment
17	cop_at11	High	Quality
18	corr_1260d	Low	Seasonality
19	coskew_21d	Low	Seasonality
20	cowc_gr1a	Low	Accruals
21	dbnetis_at	Low	Seasonality
22	debt_gr3	Low	Debt Issuance
23	debt_me	High	Value
24	div12m_me	High	Value
25	dolvol_126d	Low	Size
26	dolvol_var_126d	Low	Profitability
27	dsale_dinv	High	Profit Growth
28	ebit_bev	High	Profitability
29	ebit_sale	High	Profitability
30	ebitda_mev	High	Value
31	emp_gr1	Low	Investment
32	eq_dur	Low	Value
33	eqnetis_at	Low	Value
34	eqnpo_12m	High	Value
35	eqnpo_me	High	Value
36	eqpo_me	High	Value
37	f_score	High	Profitability
38	fcf_me	High	Value
39	fnl_gr1a	Low	Debt Issuance
40	gp_at	High	Quality
41	inv_gr1	Low	Investment
42	inv_gr1a	Low	Investment
43	iskew_ff3_21d	Low	Short-Term Reversal
44	ivol_capm_252d	Low	Low Risk

45	ivol_ff3_21d	Low	Low Risk
46	kz_index	High	Seasonality
47	lnoa_gr1a	Low	Investment
48	lti_gr1a	Low	Seasonality
49	market_equity	Low	Size
50	mispricing_mgmt	High	Investment
51	mispricing_perf	High	Quality
52	ncoa_gr1a	Low	Investment
53	netdebt_me	Low	Low Leverage
54	netis_at	Low	Value
55	nfna_gr1a	High	Debt Issuance
56	ni_be	High	Profitability
57	ni_me	High	Value
58	niq_at	High	Quality
59	niq_be	High	Profitability
60	niq_su	High	Profit Growth
61	nncoa_gr1a	Low	Investment
62	noa_at	Low	Debt Issuance
63	noa_gr1a	Low	Investment
64	o_score	Low	Profitability
65	oaccruals_at	Low	Accruals
66	oaccruals_ni	Low	Accruals
67	ocf_at	High	Profitability
68	ocf_at_chg1	High	Profit Growth
69	ocf_me	High	Value
70	ocfq_saleq_std	Low	Low Risk
71	op_at	High	Quality
72	op_atl1	High	Quality
73	ope_be	High	Profitability
74	opex_at	High	Quality
75	pi_nix	High	Seasonality
76	ppeinv_gr1a	Low	Investment
77	prc	Low	Size
78	prc_highprc_252d	High	Momentum
79	qmj	High	Quality
80	qmj_growth	High	Quality
81	qmj_prof	High	Quality
82	qmj_safety	High	Quality
83	rd_me	High	Size
84	resff3_12_1	High	Momentum
85	resff3_6_1	High	Momentum
86	ret_1_0	Low	Short-Term Reversal
87	ret_12_1	High	Momentum
88	ret_12_7	High	Profit Growth
89	ret_3_1	High	Momentum

90	ret_6_1	High	Momentum
91	ret_60_12	Low	Investment
92	ret_9_1	High	Momentum
93	rmax1_21d	Low	Low Risk
94	rmax5_21d	Low	Low Risk
95	rmax5_rvol_21d	Low	Short-Term Reversal
96	rskew_21d	Low	Short-Term Reversal
97	rvol_21d	Low	Low Risk
98	sale_bev	High	Quality
99	sale_gr1	Low	Investment
100	sale_gr3	Low	Investment
101	sale_me	High	Value
102	saleq_su	High	Profit Growth
103	seas_1_1an	High	Profit Growth
104	seas_1_1na	High	Momentum
105	seas_11_15an	High	Seasonality
106	seas_16_20an	High	Seasonality
107	seas_16_20na	Low	Accruals
108	seas_2_5an	High	Seasonality
109	seas_2_5na	Low	Investment
110	seas_6_10an	High	Seasonality
111	seas_6_10na	Low	Low Risk
112	taccruals_at	Low	Accruals
113	taccruals_ni	Low	Accruals
114	tax_gr1a	High	Profit Growth
115	turnover_126d	Low	Low Risk
116	turnover_var_126d	Low	Profitability
117	z_score	High	Low Leverage
118	zero_trades_126d	High	Low Risk
119	zero_trades_252d	High	Low Risk

Note: The table shows information about the 119 equity factors used in the paper. “Factor” is the name of the factors and the underlying factor characteristic; Long indicates whether the factor is long high or low values of the characteristic, and Theme indicates whether factors are related. The naming convention, factor direction, and cluster assignment all come from Jensen et al. (2022a). The direction ensures that the factor should have a positive realized return according to the original paper that proposed it.

Table AII. Required Factor Return

Factor	Ret	σ	$VL/s_t^{VL,i}$	VL/β_t^i	$VL/SVIX_t^i$	$IB/s_t^{VL,i}$	IB/β_t^i	$IB/SVIX_t^i$	
1	ivol_capm_252d	0.6%	18.1%	-5.5%	-2.3%	-4.7%	-10.4%	-4.6%	-8.8%
2	rvol_21d	2.0%	18.7%	-4.2%	-2.3%	-4.1%	-7.8%	-4.6%	-7.6%
3	ivol_ff3_21d	3.4%	15.3%	-4.1%	-1.7%	-3.7%	-7.6%	-3.4%	-6.9%
4	zero_trades_252d	1.3%	11.9%	-3.9%	-2.3%	-3.9%	-7.3%	-4.5%	-7.2%
5	beta_60m	-1.1%	18.2%	-3.9%	-4.1%	-3.4%	-7.3%	-8.1%	-6.2%
6	rmax5_21d	3.8%	17.5%	-3.9%	-2.2%	-3.6%	-7.3%	-4.4%	-6.7%
7	turnover_126d	0.4%	13.7%	-3.9%	-2.2%	-3.9%	-7.2%	-4.3%	-7.2%

8	div12m_me	0.9%	14.3%	-3.9%	-1.7%	-3.1%	-7.2%	-3.3%	-5.8%
9	zero_trades_126d	0.8%	12.3%	-3.8%	-2.2%	-3.8%	-7.1%	-4.3%	-7.2%
10	rmax1_21d	3.3%	15.9%	-3.5%	-1.9%	-3.3%	-6.6%	-3.8%	-6.2%
11	qmj_safety	-0.5%	15.6%	-3.5%	-1.7%	-1.2%	-6.6%	-3.4%	-2.2%
12	betabab_1260d	-0.8%	19.3%	-3.4%	-4.2%	-3.9%	-6.3%	-8.4%	-7.3%
13	eqnpo_me	5.0%	14.1%	-3.3%	-1.3%	-2.1%	-6.1%	-2.7%	-3.9%
14	eqnpo_12m	2.2%	12.8%	-3.2%	-1.2%	-2.2%	-5.9%	-2.5%	-4.0%
15	o_score	1.9%	8.2%	-3.1%	-0.3%	0.3%	-5.8%	-0.6%	0.5%
16	eqpo_me	3.9%	14.6%	-3.0%	-1.4%	-1.9%	-5.7%	-2.7%	-3.5%
17	betadown_252d	-1.1%	17.4%	-3.0%	-3.2%	-3.3%	-5.6%	-6.3%	-6.2%
18	ni_be	2.6%	9.2%	-2.6%	-0.4%	-0.7%	-5.0%	-0.8%	-1.2%
19	ebit_sale	0.4%	10.2%	-2.5%	-0.8%	-1.0%	-4.7%	-1.7%	-1.9%
20	chsho_12m	3.6%	7.7%	-2.3%	-0.6%	-1.2%	-4.4%	-1.2%	-2.3%
21	prc_highprc_252d	0.9%	18.2%	-2.3%	-1.2%	-2.3%	-4.2%	-2.4%	-4.3%
22	ni_me	4.5%	13.5%	-2.2%	-0.8%	-1.5%	-4.2%	-1.6%	-2.9%
23	eqnetis_at	5.5%	12.0%	-2.2%	-0.6%	-1.0%	-4.2%	-1.2%	-1.9%
24	niq_be	4.0%	9.4%	-2.2%	-0.3%	-0.5%	-4.1%	-0.6%	-1.0%
25	ebit_bev	3.7%	8.3%	-2.1%	-0.3%	-0.0%	-4.0%	-0.6%	-0.0%
26	qmj	3.5%	7.1%	-2.1%	-0.8%	-0.5%	-3.9%	-1.6%	-1.0%
27	ocfq_saleq_std	1.4%	8.9%	-2.1%	-1.0%	-1.4%	-3.8%	-2.0%	-2.7%
28	dolvola_var_126d	0.5%	8.1%	-2.0%	0.3%	-1.3%	-3.7%	0.6%	-2.3%
29	niq_at	3.0%	9.7%	-1.8%	-0.6%	-0.2%	-3.4%	-1.3%	-0.3%
30	ope_be	3.8%	8.8%	-1.8%	-0.3%	-0.5%	-3.4%	-0.6%	-1.0%
31	turnover_var_126d	0.4%	7.9%	-1.8%	0.4%	-1.1%	-3.4%	0.7%	-2.1%
32	mispricing_perf	6.1%	13.8%	-1.7%	-0.5%	0.0%	-3.1%	-1.1%	0.0%
33	mispricing_mgmt	5.2%	8.3%	-1.6%	-0.5%	-1.0%	-2.9%	-1.0%	-1.8%
34	fcf_me	5.1%	8.5%	-1.3%	-0.4%	-0.8%	-2.5%	-0.8%	-1.5%
35	qmj_prof	4.3%	6.9%	-1.3%	-0.3%	0.1%	-2.4%	-0.6%	0.2%
36	ocf_at	5.1%	7.6%	-1.1%	-0.5%	0.2%	-2.0%	-0.9%	0.4%
37	f_score	3.0%	7.5%	-1.0%	-0.5%	-0.7%	-1.8%	-0.9%	-1.2%
38	op_at	3.7%	7.2%	-0.9%	-0.3%	0.4%	-1.6%	-0.5%	0.8%
39	z_score	0.8%	11.0%	-0.8%	0.2%	0.8%	-1.6%	0.4%	1.5%
40	sale_gr3	0.9%	8.2%	-0.7%	-0.5%	-1.1%	-1.3%	-1.0%	-2.0%
41	op_atl1	2.7%	7.7%	-0.7%	-0.2%	0.6%	-1.3%	-0.4%	1.2%
42	emp_gr1	1.4%	8.7%	-0.6%	-0.5%	-0.8%	-1.1%	-0.9%	-1.5%
43	be_gr1a	1.5%	9.6%	-0.6%	-0.3%	-0.9%	-1.1%	-0.6%	-1.6%
44	seas_2_5na	3.3%	10.9%	-0.5%	-0.8%	-1.2%	-1.0%	-1.6%	-2.2%
45	sale_gr1	1.6%	9.8%	-0.5%	-0.3%	-0.7%	-0.9%	-0.7%	-1.4%
46	seas_6_10na	3.4%	7.3%	-0.5%	-0.9%	-1.3%	-0.9%	-1.8%	-2.4%
47	ppeinv_gr1a	4.3%	7.4%	-0.4%	-0.1%	-0.4%	-0.8%	-0.2%	-0.7%
48	capx_gr3	2.3%	6.9%	-0.4%	-0.3%	-0.6%	-0.8%	-0.6%	-1.2%
49	rskew_21d	2.2%	4.5%	-0.4%	-0.2%	-0.3%	-0.8%	-0.4%	-0.5%
50	seas_16_20na	0.6%	5.6%	-0.4%	-0.3%	-0.4%	-0.7%	-0.6%	-0.7%
51	gp_at	3.8%	7.9%	-0.4%	-0.2%	0.5%	-0.7%	-0.3%	0.9%
52	iskew_ff3_21d	0.3%	3.4%	-0.4%	-0.2%	-0.1%	-0.7%	-0.3%	-0.2%

53	noa_at	5.2%	5.4%	-0.4%	0.2%	0.5%	-0.7%	0.4%	0.9%
54	eq_dur	3.4%	11.8%	-0.4%	-0.4%	-0.9%	-0.7%	-0.8%	-1.6%
55	capx_gr2	3.4%	6.7%	-0.3%	-0.2%	-0.5%	-0.7%	-0.4%	-1.0%
56	ebitda_mev	5.7%	13.2%	-0.3%	-0.3%	-0.6%	-0.6%	-0.6%	-1.1%
57	netdebt_me	1.1%	10.9%	-0.3%	0.6%	1.4%	-0.5%	1.1%	2.6%
58	ret_12_1	7.2%	16.4%	-0.2%	-0.2%	-0.4%	-0.4%	-0.3%	-0.7%
59	netis_at	4.9%	9.9%	-0.2%	-0.1%	-0.2%	-0.4%	-0.3%	-0.4%
60	inv_gr1	4.5%	6.3%	-0.2%	-0.2%	-0.4%	-0.4%	-0.4%	-0.8%
61	qmj_growth	3.0%	5.5%	-0.2%	0.0%	-0.3%	-0.4%	0.1%	-0.6%
62	cop_at11	5.4%	6.1%	-0.2%	0.1%	0.8%	-0.3%	0.1%	1.5%
63	col_gr1a	-0.3%	8.2%	-0.1%	-0.3%	-0.7%	-0.3%	-0.7%	-1.3%
64	resff3_12_1	6.9%	7.6%	-0.1%	-0.3%	-0.4%	-0.2%	-0.5%	-0.7%
65	resff3_6_1	3.6%	7.1%	-0.1%	-0.2%	-0.4%	-0.2%	-0.4%	-0.8%
66	ret_12_7	6.6%	13.6%	-0.1%	-0.0%	-0.0%	-0.2%	-0.1%	-0.0%
67	ret_9_1	5.5%	16.1%	-0.1%	-0.1%	-0.3%	-0.2%	-0.2%	-0.6%
68	lti_gr1a	0.3%	3.6%	-0.1%	-0.0%	0.1%	-0.2%	-0.1%	0.1%
69	coa_gr1a	2.6%	7.7%	-0.1%	-0.4%	-0.5%	-0.1%	-0.8%	-1.0%
70	ocf_me	5.7%	12.2%	-0.1%	-0.3%	-0.6%	-0.1%	-0.5%	-1.2%
71	ret_1_0	6.1%	12.2%	-0.0%	-0.0%	0.2%	-0.1%	-0.0%	0.4%
72	noa_gr1a	5.5%	7.0%	-0.0%	-0.2%	-0.3%	-0.1%	-0.3%	-0.6%
73	seas_1_1an	4.2%	10.1%	-0.0%	-0.0%	0.0%	-0.0%	-0.0%	0.1%
74	inv_gr1a	3.7%	5.8%	-0.0%	-0.2%	-0.3%	-0.0%	-0.4%	-0.6%
75	ret_6_1	3.0%	15.8%	-0.0%	-0.0%	-0.3%	-0.0%	-0.1%	-0.5%
76	lnoa_gr1a	4.0%	6.0%	-0.0%	0.1%	-0.4%	-0.0%	0.1%	-0.7%
77	cowc_gr1a	3.9%	4.7%	-0.0%	-0.2%	0.1%	-0.0%	-0.5%	0.1%
78	seas_2_5an	3.8%	8.3%	0.0%	0.2%	0.3%	0.1%	0.4%	0.6%
79	seas_16_20an	2.8%	5.4%	0.0%	0.0%	0.1%	0.1%	0.1%	0.3%
80	niq_su	2.9%	5.2%	0.0%	0.0%	0.0%	0.1%	0.0%	0.1%
81	seas_6_10an	4.5%	6.5%	0.0%	0.2%	0.4%	0.1%	0.4%	0.7%
82	at_gr1	2.7%	9.2%	0.1%	-0.3%	-0.7%	0.1%	-0.6%	-1.3%
83	sale_bev	4.2%	6.9%	0.1%	0.3%	0.8%	0.1%	0.6%	1.5%
84	seas_11_15an	2.7%	5.1%	0.1%	0.1%	0.1%	0.1%	0.2%	0.2%
85	ret_3_1	1.3%	13.4%	0.1%	0.0%	-0.2%	0.1%	0.1%	-0.3%
86	tax_gr1a	1.4%	6.4%	0.1%	0.1%	0.4%	0.2%	0.3%	0.7%
87	nncoa_gr1a	4.6%	5.8%	0.1%	0.0%	-0.3%	0.2%	0.0%	-0.6%
88	ocf_at_chg1	2.3%	3.7%	0.1%	0.1%	0.0%	0.3%	0.1%	0.0%
89	dsale_dinv	3.3%	3.9%	0.2%	0.0%	0.1%	0.3%	0.0%	0.1%
90	coskew_21d	0.4%	5.6%	0.2%	-0.1%	-0.1%	0.3%	-0.2%	-0.1%
91	opex_at	3.0%	7.3%	0.2%	0.1%	0.5%	0.4%	0.1%	0.9%
92	saleq_su	1.2%	6.2%	0.3%	0.0%	0.1%	0.5%	0.1%	0.2%
93	pi_nix	1.1%	4.6%	0.3%	-0.1%	0.1%	0.5%	-0.3%	0.1%
94	rmax5_rvol_21d	5.1%	7.4%	0.3%	0.1%	0.1%	0.6%	0.2%	0.2%
95	ncoa_gr1a	4.5%	5.6%	0.4%	0.1%	-0.2%	0.8%	0.3%	-0.3%
96	capex_abn	2.5%	4.2%	0.4%	0.2%	0.1%	0.8%	0.4%	0.3%
97	nfna_gr1a	3.0%	3.5%	0.5%	0.2%	0.2%	0.8%	0.3%	0.4%

98	taccruals_at	2.2%	5.8%	0.5%	0.1%	0.4%	0.9%	0.2%	0.8%
99	seas_1_1na	3.6%	16.1%	0.5%	0.2%	0.0%	0.9%	0.4%	0.1%
100	fnl_gr1a	2.9%	3.5%	0.6%	0.2%	0.2%	1.0%	0.4%	0.3%
101	debt_gr3	2.5%	4.0%	0.6%	0.2%	-0.1%	1.0%	0.4%	-0.1%
102	be_me	2.5%	11.9%	0.7%	-0.2%	-0.3%	1.3%	-0.4%	-0.5%
103	taccruals_ni	1.9%	4.8%	0.8%	0.1%	0.2%	1.4%	0.2%	0.5%
104	debt_me	0.3%	12.4%	0.8%	-0.1%	-1.0%	1.4%	-0.2%	-1.8%
105	bev_mev	1.7%	12.6%	0.9%	0.1%	-0.6%	1.7%	0.1%	-1.0%
106	dbnetis_at	2.2%	4.1%	1.1%	0.2%	0.1%	2.0%	0.5%	0.2%
107	corr_1260d	0.3%	10.9%	1.1%	-3.0%	-0.8%	2.1%	-5.8%	-1.6%
108	oaccruals_at	4.8%	6.1%	1.1%	0.2%	0.8%	2.1%	0.4%	1.4%
109	sale_me	4.4%	12.4%	1.2%	0.2%	-0.1%	2.2%	0.5%	-0.2%
110	rd_me	4.8%	10.0%	1.2%	0.9%	1.0%	2.3%	1.9%	1.9%
111	ret_60_12	2.5%	11.5%	1.4%	0.2%	-0.1%	2.6%	0.4%	-0.2%
112	dolvol_126d	1.7%	9.5%	1.5%	-1.1%	-1.2%	2.8%	-2.2%	-2.2%
113	oaccruals_ni	4.2%	4.9%	1.5%	0.3%	0.6%	2.8%	0.6%	1.1%
114	kz_index	-1.0%	6.8%	2.0%	-0.2%	0.2%	3.8%	-0.3%	0.4%
115	ami_126d	1.9%	11.1%	2.6%	-0.5%	0.5%	4.9%	-0.9%	1.0%
116	age	-1.0%	10.5%	2.9%	1.2%	2.2%	5.5%	2.3%	4.1%
117	market_equity	2.4%	12.6%	3.6%	-0.0%	1.2%	6.8%	-0.0%	2.2%
118	bidaskhl_21d	-0.6%	15.1%	3.8%	1.9%	3.3%	7.2%	3.7%	6.2%
119	prc	1.7%	15.0%	4.0%	1.0%	1.6%	7.6%	1.9%	3.1%

Note: The table shows the required and realized return of the factors from Table AI. “Ret” is the annualized average realized return of the factor and “ σ ” is the annualized volatility of the realized return. The remaining columns show the annualized required factor return for the six required return proxies. The name before the slash indicates whether the subjective expected return proxy is from Value Line (VL) or I/B/E/S (IB). The name after the slash shows the subjective risk proxy which is either the safety rank from Value Line ($s_t^{VL,i}$), the market beta estimate from Value Line (β_t^i), or the $SVIX_t^i$ measure from Martin and Wagner (2019). A bold number indicates that the p -value from (1.33) is below 5% meaning that realized return is significantly different from the required return i.e. that the “risk hypothesis” is rejected. The factors are sorted according to the required return in the “VL/ $s_t^{VL,i}$ ” column.

8.8 Machine Learning Predictions via XGBoost

In this section, I describe the approach I use to make machine learning predictions for the ML model from Section 6. I use the XGBoost model from Chen and Guestrin (2016), with the 119 characteristics from Table AI as inputs, and the outcome variable is a stock’s realized excess return over the next month.

The first model is based on training data from 1952 to 1971. However, XGBoost requires me to decide on several “hyper-parameters” such as the number of decision trees to use in the ensemble and the maximum tree depth. To choose these hyper-parameters, I use the last 10 years of the training period as the validation period. For a set of hyper-parameters, I train the model on the data prior to the validation period and record its



Figure A3. The Subjective Risk Premium is Countercyclical

Note: The top-left panel shows the annualized required market return where required returns are estimated from (1.21) using the safety rank and subjective expected return from Value Line. The top-right panel shows an index of the relative risk aversion from Bekaert et al. (2022), the lower-left panel shows the one-year expected market return from the Livingston survey of professional forecasters, and the lower-right panel shows the one-year expected market return of retail investors from Nagel and Xu (2022a).

mean squared error (MSE) on the validation data. I repeat this for 20 sets of hyper-parameters shown in Table AIII and choose the set of hyper-parameters with the lowest MSE. Using these hyper-parameters, I then re-train the model on the full training data giving me the first model, $\hat{f}_1(x_t^i)$, where $x_t^i \in \mathbb{R}^{119 \times 1}$ is the vector of stock characteristics. I then use $\hat{f}_1(x_t^i)$ to predict returns from 1972-1981.

I update the model each decade by expanding the training, validation, and test period by 10 years. For example, the second model is based on 1952-1981 as the training period, 1972-1981 as the validation period, and 1982-1991 as the test period. I fit five different models that predict returns out-of-sample from 1972 to 2021.

8.9 Bootstrap p -Values for Comparing Asset Pricing Models

This section shows how I compute the “bootstrap p -value” used for the pairwise model comparison test in Section 6. My approach is inspired by chapter 11 in Efron and Hastie (2016). I first describe the bootstrap sampling procedure and then how I compute the confidence distribution based on these bootstrap estimates.

Table AIII. XGBoost Hyper-parameters

	No. Features	Tree depth	Learning rate	Sample size	Penalty
1	103	3	0.017	0.20	0.05
2	105	3	0.011	0.91	2.33
3	26	1	0.222	0.99	0.02
4	12	1	0.046	0.27	52.44
5	16	2	0.012	0.93	48.81
6	49	1	0.153	0.38	1.55
7	28	2	0.293	0.65	72.28
8	18	2	0.059	0.87	33.11
9	83	1	0.103	0.95	77.43
10	32	3	0.224	0.40	9.85
11	88	2	0.110	0.36	0.21
12	95	2	0.024	0.71	33.64
13	103	2	0.012	0.57	2.49
14	80	3	0.067	0.44	0.19
15	40	2	0.061	0.64	2.31
16	118	2	0.202	0.64	0.01
17	7	2	0.011	0.54	0.02
18	81	1	0.108	0.72	0.08
19	94	2	0.292	0.75	2.27
20	66	2	0.026	0.71	0.06

Note: The table shows the hyper-parameters considered for the XGBoost model used in Section 6. “No. Features” is the number of randomly chosen features considered for each decision tree, “Tree depth” is the maximum depth of each decision tree, “Learning rate” is the weight each new tree gets in the ensemble, “Sample size” is the fraction of the observations randomly chosen to train each decision tree on, and “Penalty” is an L2 (ridge) penalty. I get the hyper-parameter sets by specifying a tolerable range for each hyper-parameter and then use the `grid_max_entropy` function from the `dials` package (<https://dials.tidymodels.org/>) to get 20 sets that aim to cover the associated parameter space.

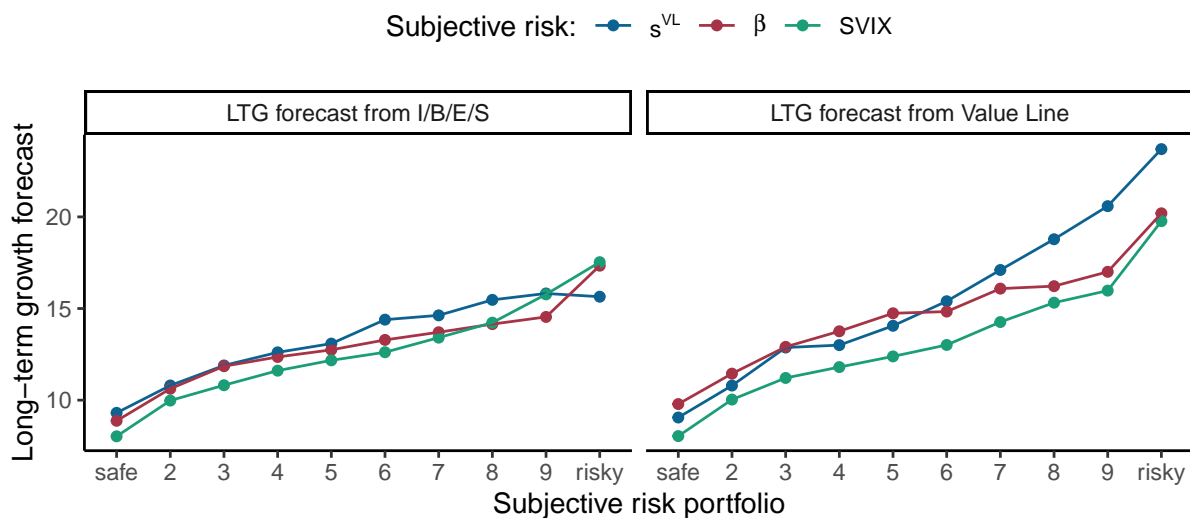


Figure A4. Riskier stocks have higher expected long-term earnings growth

Note: The figure shows the expected long-term growth (LTG) in earnings per share (EPS) for portfolios sorted on subjective risk. Each month, I sort stocks into ten portfolio based on their subjective risk, compute the average LTG within each portfolio, and average these numbers over time. In the left panel, LTG is the median consensus long-term growth forecast from sell-side analysts available from I/B/E/S, defined as the expected annual increase in EPS over the company’s next full business cycle. In the right panel, LTG is a forecast of the annualized growth in EPS over the next three-to-five years from Value Line.

As in Section 5, I need to account for the high persistence of required returns. In addition, I also need to account for the strong correlation of realized and required returns across assets. Therefore, when realized and required returns overlap, I use a moving block bootstrap procedure with data from $n = 30$ consecutive months if the subjective expected return proxy is from Value Line and $n = 12$ if it is from I/B/E/S. In the non-overlapping period, I only have monthly return data, which generally have a low autocorrelation but substantial cross-sectional correlation. Therefore, I sample data from individual months rather than temporal blocks. Let T_1 and T_2 be the number of months in the non-overlapping and overlapping period, respectively. To compute a full bootstrap sample, I start by randomly choosing data from T_1 months from the non-overlapping period with replacements, then I randomly choose $\lceil T_2/n \rceil$ blocks from the non-overlapping period but delete the last $\lceil T_2/n \rceil n - T_2$ months (to ensure that the sample length is T_2); finally, I combine these two sub-samples. For each bootstrap sample, b , I compute R_b^2 and \tilde{R}_b^2 for each model and each required return proxy. I repeat this procedure $B = 2,500$ times. Importantly, sample data from a specific month includes portfolio returns, pricing factor returns, consumption data, subjective risk, and subjective return expectations. This procedure ensures that I account for all the sampling variability related to the estimation.

Next, I describe the confidence distribution. I focus on the realized pricing ability,

but the procedure is identical for the required pricing ability. Let $\hat{d} = R_{m_1}^2 - R_{m_2}^2$ be the difference in realized pricing ability of model m_1 and m_2 . Having a confidence distribution of \hat{d} , I am interested in $h(0)$. A low value shows support for m_1 while a high value shows support for m_2 . Let \hat{d}^b be the pricing difference in bootstrap sample b . We can create a simple bootstrap confidence distribution of \hat{d}_{m_1, m_2} using the percentile method (Efron and Hastie, 2016). Here, the confidence distribution is simply the proportion of the bootstrap estimates that are below a specific value x :

$$h_{perc}(x) = \frac{\#\{\hat{d}^b \leq x\}}{B}. \quad (1.56)$$

Here, $h_{perc}(0)$ is simply the proportion of bootstrap estimates below zero. It is useful to have the relation in (1.56) in mind when interpreting the confidence distribution. However, the standard percentile method can be biased, so I instead use the bias-corrected percentile method (Efron and Hastie, 2016). Let $c = h_{perc}(\hat{d})$ be the proportion of bootstrap estimates smaller than or equal to the data estimate. If $c \neq 0.5$, the bootstrap distribution is median biased. Next, let the bias correction value be $z = \Phi^{-1}(c)$, where $\Phi^{-1}(x)$ is the inverse of the standard normal distribution. The bias-corrected inverse distribution is

$$h_{BC}^{-1}(\alpha) = h_{perc}^{-1}(\Phi[2z + \Phi^{-1}(\alpha)]), \quad (1.57)$$

where h_{perc}^{-1} is the inverse of (1.56) and $\Phi(x)$ is the standard normal distribution. The standard and bias-corrected percentile function coincide if and only if the bootstrap distribution is median unbiased, $z = 0$: $h_{BC}^{-1}(\alpha) = h^{-1}(\Phi[\Phi^{-1}(\alpha)]) = h^{-1}(\alpha)$. I invert (1.57) to obtain my final confidence distribution, $h_{BC}(x)$. The bootstrap p -value is the value of this distribution at $x = 0$.

8.10 Robustness

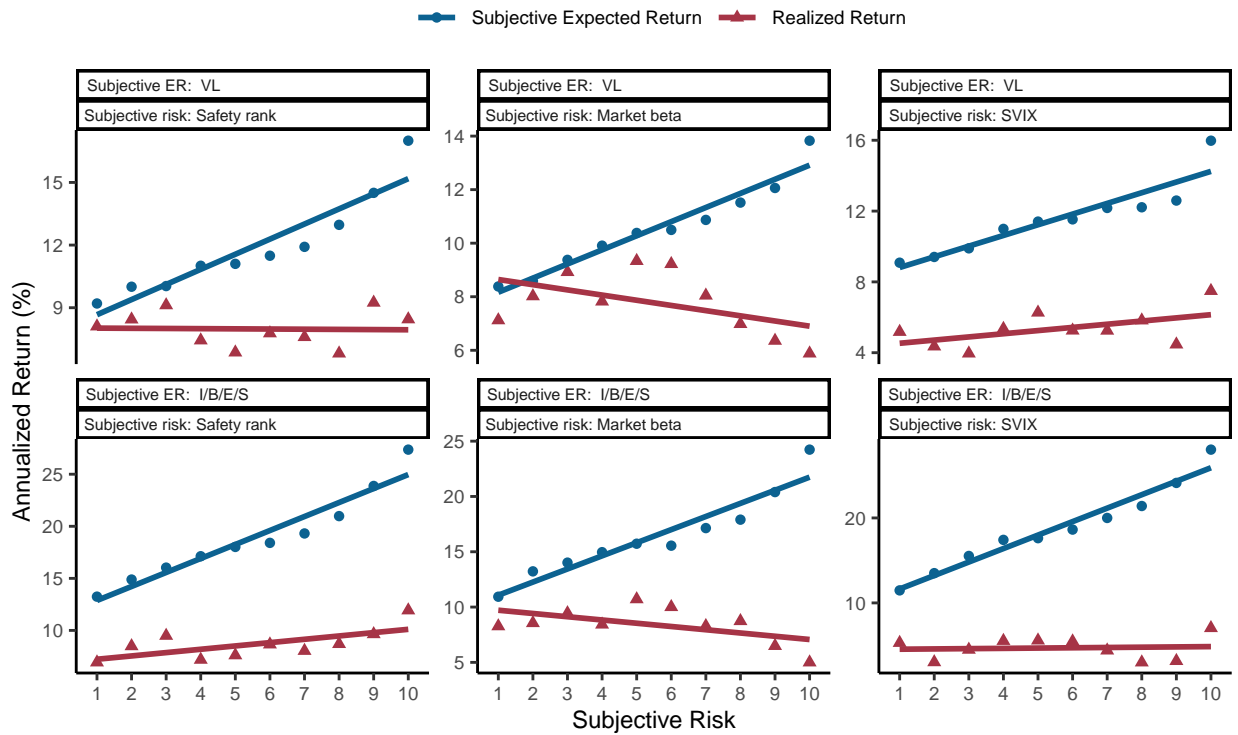


Figure A5. The Subjective and Realized Risk-Return Tradeoff—Value Weights

Note: This figure is identical to Figure 1 except that stocks are aggregated into portfolio using value-weights instead of equal weights.

Table AIV. Subjective Risk and Objective Mispricing—Alternative Proxies

	$(\tilde{E}_t^{VL}[\text{eps}_{t+4}^i] - \text{eps}_{t+4}^i)/p_t^i \times 100$			$\tilde{E}_t^{VL}[\text{pe}_{t+4}^i] - \text{pe}_{t+4}^i$		
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	1.25 (4.91)	1.69 (5.88)	1.14 (3.15)	-4.10 (-6.82)	-4.04 (-6.68)	-1.06 (-1.30)
$s_t^{VL,i}$	4.83 (12.45)			5.65 (5.83)		
β_t^i		3.58 (8.88)			5.04 (4.82)	
SVIX_t^i			3.19 (5.99)			1.33 (1.03)
Observations	26,345	26,345	5,708	25,844	25,844	5,663
Adjusted R ²	0.03	0.02	0.02	0.00	0.00	0.00

Note: The table shows estimates from a regression of forecast error on subjective risk. It supplements Table V with alternative forecast variables. In columns 1-3 forecast error is defined as the forecast of earnings per share (EPS) in the fiscal year 4 years from the forecast date minus the subsequent realization scaled by the price issued at the beginning of the forecast month. In columns 4-6 forecast error is defined as the forecast of the price-to-earnings ratio in the fiscal year 4 years from the forecast date minus the subsequent realization. For each firm-fiscal year I only retain the first forecast issued at least 45 days and at most 180 days after the announcement of the most recent fiscal year's EPS. All forecast are from Value Line. The dependent variables are winsorized at the top/bottom 1%. The subjective risk proxies are scaled to lie between 0 and 1 each month, so the intercept shows the forecast error on the safest stocks (risk=0), and the slope coefficient shows the forecast error from the riskiest stock (risk=1). The number in the parenthesis shows the t -statistic based on standard errors clustered by firm and quarter of the fiscal year-end.

Chapter 2

Is there a Replication Crisis in Finance?

with Bryan Kelly and Lasse Heje Pedersen.

Abstract

Several papers argue that financial economics faces a replication crisis because the majority of studies cannot be replicated or are the result of multiple testing of too many factors. We develop and estimate a Bayesian model of factor replication, which leads to different conclusions. The majority of asset pricing factors: (1) can be replicated, (2) can be clustered into 13 themes, the majority of which are significant parts of the tangency portfolio, (3) work out-of-sample in a new large data set covering 93 countries, and (4) have evidence that is strengthened (not weakened) by the large number of observed factors.

Kelly is at Yale School of Management, AQR Capital Management, and NBER. Pedersen is at AQR Capital Management, Copenhagen Business School, and CEPR. We are grateful for helpful comments from Nick Barberis, Andrea Frazzini, Cam Harvey (discussant), Antti Ilmanen, Ronen Israel, Andrew Karolyi, John Liew, Toby Moskowitz, Stefan Nagel, Scott Richardson, Anders Rønn-Nielsen, Neil Shephard (discussant), and seminar and conference participants at AFA 2022, NBER 2021, AQR, Georgetown Virtual Fintech Seminar, Tisvildeleje Summer Workshop 2020, Yale, and the CFA Institute European Investment Conference 2020. We thank Tyler Gwinn for excellent research assistance. Jensen and Pedersen gratefully acknowledge support from the FRIC Center for Financial Frictions (grant no. DNRF102). AQR Capital Management is a global investment management firm, which may or may not apply similar investment techniques or methods of analysis as described herein. The views expressed here are those of the authors and not necessarily those of AQR.

Several research fields face replication crises (or credibility crises), including medicine (Ioannidis, 2005), psychology (Nosek et al., 2012), management (Bettis, 2012), experimental economics (Maniadis et al., 2017), and now also financial economics. Challenges to the replicability of finance research take two basic forms:

1. **No internal validity.** Most studies cannot be replicated with the same data (e.g., because of coding errors or faulty statistics) or are not robust in the sense that the main results cannot be replicated using slightly different methodologies and/or slightly different data.¹ E.g., Hou et al. (2020b) state:

“Most anomalies fail to hold up to currently acceptable standards for empirical finance.”

2. **No external validity.** Most studies may be robustly replicated, but are spurious and driven by “*p*-hacking,” that is, finding significant results by testing multiple hypotheses without controlling the false discovery rate. Such spurious results are not expected to replicate in other samples or time periods, in part because the sheer number of factors is simply too large, and too fast growing, to be believable. E.g., Cochrane (2011) asks for a consolidation of the “factor zoo,” and Harvey et al. (2016b) states:

*“most claimed research findings in financial economics are likely false.”*²

We examine both of these challenges theoretically and empirically. We conclude that neither criticism is tenable. The majority of factors do replicate, do survive joint modeling of all factors, do hold up out-of-sample, are strengthened (not weakened) by the large number of observed factors, are further strengthened by global evidence, and the number of factors can be understood as multiple versions of a smaller number of themes.

These conclusions rely on new theory and data: First, we show that factors must be understood in light of economic theory and we develop a Bayesian model that offers a very different interpretation of the evidence on factor replication. Second, we put together a new global data set of 153 factors across 93 countries. To help advance replication in

¹Hamermesh (2007) contrasts “pure replication” and “scientific replication.” Pure replication is, “checking on others’ published papers using their data,” also called “reproduction” by Welch (2019). Scientific replication uses, “different sample, different population and perhaps similar, but not identical model.” We focus on scientific replication. We propose a new modeling framework to jointly estimate factor alphas, we use robust factor construction methods that are applied uniformly to all factors, and we test both internal and external validity of prior factor research in several dimensions, including out-of-sample time series replication and international sample replication. In complementary and contemporaneous work, Chen and Zimmermann (2020a) consider pure replication, attempting to use the same data and methods as the original papers for a large number of factors. They are able to reproduce nearly 100% of factors, but Hou et al. (2020b) challenge the scientific replication and Harvey et al. (2016b) challenge validity due to multiple testing.

²Similarly, Linnainmaa and Roberts (2018) state “the majority of accounting-based return anomalies, including investment, are most likely an artifact of data snooping.”

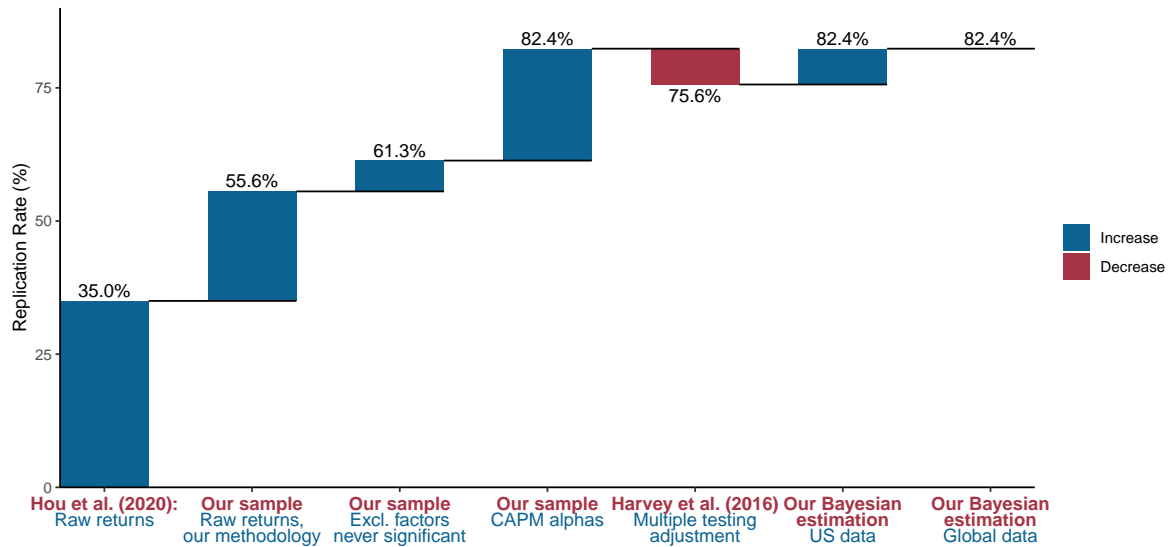


Figure 1. Replication Rates Versus the Literature

Note: This figure summarizes analyses throughout the paper. Refer to Section 3 for estimation details.

finance, we have made this data set easily accessible to researchers via a direct open-source link to WRDS, including meticulous documentation of the data and the underlying code base.

Replication results. Figure 1 illustrates our main results and how they relate to the literature in a sequence of steps. It presents the “replication rate,” that is, the percent of factors with a statistically significant average excess return. The starting point of Figure 1—shown as the first bar on the left—is the 35% replication rate reported in the expansive factor replication study of Hou et al. (2020b).

The second bar in Figure 1 shows a 55.6% baseline replication rate in our main sample of US factors. It is based on significant OLS t -statistics for average raw factor returns, in direct comparability to the 35% calculation from Hou et al. (2020b). This difference arises because our sample is longer, we add 15 factors to our sample that were previously studied in the literature but not studied by Hou et al. (2020b), and due to minor conservative factor construction details that we believe robustify factor behavior.³ We discuss this decomposition further in Section 2, where we detail our factor construction choices and discuss why we prefer them.

³We use tercile spreads while they use deciles; we use tercile breakpoints from all stocks above the NYSE 20th percentile (i.e., non-micro-caps), they use straight NYSE breakpoints; we always lag accounting data four months, they use a mixture of updating schemes; we exclude IBES factor due to their relatively short history; we use capped value-weighting they use straight value-weights; We look at returns over a 1 month holding period, they use 1, 6 and 12 months. In appendix 5.4 we detail how each change affects the replication rate.

The Hou et al. (2020b) sample includes a number of factors that the original studies found to be insignificant.⁴ We exclude these when calculating the replication rate. After we make this adjustment, the replication rate rises to 61.3%, shown in the third bar in Figure 1.

Alpha, not raw return. Hou et al. (2020b) analyze and test factors’ raw returns, but if we wish to learn about “anomalies,” economic theory dictates the use of risk-adjusted returns. Raw return gives a misleading inference for the factor if it differs from the alpha: When the raw return is significant, but the alpha is not, it simply means that the factor is taking risk exposure and the risk premium is significant, which does not indicate anomalous returns of the factor. Likewise, when the raw return is insignificant, but the alpha is significant, then the factor’s efficacy is masked by its risk exposure. An example of this is the low-beta anomaly, where theory predicts that the alpha of a dollar-neutral low-beta factor is positive, but its raw return is negative or close to zero (Frazzini and Pedersen, 2014). In this case, the “failure to replicate” of Hou et al. (2020b) is, in fact, support for the betting-against-beta theory. We analyze alpha to the CAPM, which is the clearest theoretical benchmark model that is not mechanically linked to other so-called anomalies in the list of replicated factors. The fourth bar in Figure 1 shows that the replication rate rises to 82.4% based on tests of factors’ CAPM alpha.

Multiple testing and our Bayesian model. The first four bars in Figure 1 are based on individual ordinary least squares (OLS) t -statistics for each factor. But Harvey et al. (2016b) rightly point out that this type of analysis suffers from a multiple testing (MT) problem. Harvey et al. (2016b) recommend MT adjustments that raise the threshold for a t -statistic to be considered statistically significant. We report one such MT correction using a leading method proposed by Benjamini and Yekutieli (2001). Accounting for MT in this manner, we find that the replication rate drops to 75.6% (the fifth bar of Figure 1). For comparison, Hou et al. (2020b) consider a similar adjustment and find that their replication rate drops from 35% with OLS to 18% after MT correction.

However, common frequentist MT corrections can be unnecessarily crude. Our handling of the MT problem is different. We propose a Bayesian framework for the joint behavior of all the factors, resulting in an MT correction that sacrifices much less power than its frequentist counterpart (which we demonstrate via simulation).⁵ To understand the benefits of our approach, note first that we impose a prior that all alphas are expected

⁴We identify 34 factors from Hou et al. (2020b) for which the original paper did not find a significant alpha or did not study factor returns (see appendix Table AIII).

⁵A large statistics literature (see Gelman et al., 2013, and references therein) explains how Bayesian estimation naturally combats MT problems and Gelman et al. (2012) conclude that “the problem of multiple comparisons can disappear entirely when viewed from a hierarchical Bayesian perspective.” Chincio et al. (2020) use a Bayesian estimation framework similar to ours for a different (but conceptually related) problem. They infer the distribution of coefficients in a stock return prediction model to calculate what they dub the “anomaly base rate.”

to be zero. The role of the Bayesian prior is conceptually similar to that of frequentist MT corrections—it imposes conservatism on statistical inference and controls the false discovery rate. Second, our *joint* factor model allows us to conduct inference for all factor alphas simultaneously. The joint structure among factors leverages dependence in the data in order to draw more informative statistical inferences (relative to conducting independent individual tests). Our zero-alpha prior shrinks alpha estimates of all factors, thereby leading to fewer discoveries (i.e., a lower replication rate), with similar conservatism as a frequentist MT correction. At the same, however, the model allows us to learn more about the alpha of any individual factor, borrowing estimation strength across all factors. The improved precision of alpha estimates for all factors can increase the number of discoveries. Which effect dominates when we construct our final Bayesian model—the conservative shrinkage to the prior or the improved precision of alphas—is an empirical question.

In our sample, we find that the two effects exactly offset, which is why the Bayesian multiple testing view delivers a replication rate identical to the OLS-based rate. Specifically, our estimated replication rate rises to 82.4% (the sixth bar of Figure 1) using our Bayesian approach to the MT problem.⁶ The intuition behind this surprising result is simply that having many factors (a “factor zoo”) can be a strength rather than a weakness when assessing the replicability of factor research. It is obvious that our posterior is tighter when a factor has performed better and has a longer time series. But the posterior is further tightened if similar factors have also performed well, and if additional data shows that these factors have performed well in many other countries.⁷

Benefits of our model beyond the replication rate. One of the key benefits of Bayesian statistics is that one recovers not just a point estimate but the entire posterior distribution of parameters. The posterior allows us to make any possible probability calculation about parameters. For example, in addition to the replication rate, we also calculate the posterior probability of false discoveries (false discovery rate, FDR) and the posterior expected fraction of true factors. Moreover, we calculate Bayesian confidence intervals (also called credibility intervals) for each of these estimates. We find that our 82.4% replication rate has a tight posterior standard error of 2.8%. The posterior Bayesian FDR is only 0.1% with a 95% confidence interval of [0.0%, 1.0%], demonstrating the small risk of false discoveries. The expected fraction of true factors is 94.0% with a posterior

⁶Our Bayesian approach leads to an even larger increase in the replication rate when using pure value-weighted returns (see Figure A1 of the appendix) and when considering global evidence outside the US (as we show later, in Figure 6).

⁷Taking this intuition further, we can glean additional information from studying whether factors work in other asset classes, as has been done for value and momentum (Asness et al., 2013), betting against beta (Frazzini and Pedersen, 2014), time series momentum (Moskowitz et al., 2012), and carry (Kojien et al., 2018).

standard error of 1.3%.

Global replication. Having found a high degree of internal validity of prior research, we next consider external validity across countries and over time. Regarding the former, we investigate how our conclusions are affected when we extend the data to include all factors in a large global panel of 93 countries. The last bar in Figure 1, shows that, based on the global sample, the final replication rate is 82.4%. This estimate is based on the Bayesian model extended to incorporate the joint behavior of international data. Because it accounts for the global correlation structure among factors, the model recognizes that international evidence is not independent out-of-sample evidence, and uses only the incremental global evidence to update the overall replicability assessment. And it continues to account for multiple testing. The global result reflects that factor performance in the US replicates well in an extensive cross section of countries. Serving as our final estimate, the global factor replication rate more than doubles that of Hou et al. (2020b) by grounding our tests in economic theory and modern Bayesian statistics. We conclude from the global analysis that factor research demonstrates external validity in the cross section of countries.

Post-publication performance. McLean and Pontiff (2016) find that US factor returns “are 26% lower out-of-sample and 58% lower post-publication.”⁸ Our Bayesian framework shows that, given a prior belief of zero alpha but an OLS alpha ($\hat{\alpha}$) that is positive, then our posterior belief about alpha lies somewhere between zero and $\hat{\alpha}$. Hence, a positive but attenuated post-publication alpha is the expected outcome based on Bayesian learning, rather than a sign of non-reproducibility. Further, when comparing factors cross-sectionally, the prediction of the Bayesian framework is that higher pre-publication alphas, if real, should be associated with higher post-publication alphas on average. And that is what we find. We present new and significant cross-sectional evidence that factors with higher in-sample alpha generally have higher out-of-sample alpha. The attenuation in the data is somewhat stronger than predicted by our Bayesian model. We conclude that factor research demonstrates external validity in the time series, but there appears to be some decay of the strongest factors that could be due to arbitrage or data mining.⁹

Publication bias. We also address the issue that factors with strong in-sample performance are more likely to be published while poorly performing factors are more

⁸Extending the evidence to global stock markets, Jacobs and Muller (2020) find that “the United States is the only country with a reliable post-publication decline in long-short returns.” Chen and Zimmermann (2020b) use Bayesian methods to estimate bias-corrected post-publication performance and find that average returns drop by only 12% after publication in US data.

⁹Data prior to the sample used in original studies also constitutes out-of-sample evidence (Linnainmaa and Roberts, 2018; Ilmanen et al., 2021). Our external validity conclusions are the same when we also include pre-original-study out-of-sample evidence.

likely to be unobserved in the literature. Publication bias can influence our full-sample Bayesian evidence through the empirical Bayes estimation of prior hyperparameters. To account for this bias, we show how to pick a prior distribution that is unaffected by publication bias by using only out-of-sample data or estimates from Harvey et al. (2016b). Using such priors, the full-sample alphas are shrunk more heavily toward zero. The result is a slight drop in US the replication rate to 81.5%. If we add an extra degree of conservatism to the prior, the replication rate drops to 79.8%. Further, our out-of-sample evidence across time and across countries is not subject to publication bias.

Multidimensional challenge: A Darwinian view of the factor zoo. Harvey et al. (2016b) challenge the sheer number of factors and Cochrane (2011) refers to as “the multidimensional challenge.” We argue that the factor research universe should not be viewed as hundreds of distinct factors. Instead, factors cluster into a relatively small number of highly correlated themes, and this property features prominently in our Bayesian modeling approach. We propose a factor taxonomy that algorithmically classifies factors into 13 themes possessing a high degree of within-theme return correlation and economic concept similarity, and low across-theme correlation. The emergence of themes, in which factors are minor variations on a related idea, is intuitive. For example, each value factor is defined by a specific valuation ratio, but there are many plausible ratios. Considering their variations is not spurious alpha-hacking, particularly when the “correct” value signal construction is debatable.

We estimate a replication rate of greater than 50% in 11 of the 13 themes (based on the Bayesian model including MT adjustment), the exceptions being “low leverage,” and “size” factor themes. We also analyze which themes matter when simultaneously controlling for all other themes. To do so, we estimate the ex post tangency portfolio of 13 theme-representative portfolios. We find that 10 of the 13 themes enter into the tangency portfolio with significantly positive weights, where the three displaced themes are “profitability,” “investment,” and “size.”

Why, the profession asks, have we arrived at a “factor zoo”?¹⁰ The answer, evidently, is because the risk-return tradeoff is complex and difficult to measure. The complexity manifests in our inability to isolate a single, silver bullet characteristic that pins down the risk-return tradeoff. Classifying factors into themes, we trace the economic culprits to roughly a dozen concepts. This is already a multidimensional challenge, but it is compounded by the fact that within a theme there are many detailed choices for how to configure the economic concept, which results in highly correlated within-theme factors.

¹⁰See Bryzgalova et al. (2019), Chordia et al. (2020), Kelly et al. (2019), Kozak et al. (2020), Green et al. (2017), and Feng et al. (2020) for other perspectives on high-dimensional asset pricing problems, and Chen (2020) for an argument why p -hacking cannot explain the existence of so many significant factors.

Together, the themes (and the factors in them) each make slightly different contributions to our collective understanding of markets. A more positive take on the factor zoo is *not* as a collective exercise in data mining and false discovery; instead, it is a natural outcome of a decentralized effort in which researchers make contributions that are correlated with, but incrementally improve on, the shared body of knowledge.

Economic implications. Our findings have broad implications for finance researchers and practitioners. We confirm that the body of finance research contains a multitude of replicable information about the drivers of expected returns. Further, we show that investors would have profited from factors deemed significant by our Bayesian method, but deemed insignificant by the frequentist MT method proposed by Harvey et al. (2016b). Indeed, Figure 2 plots the returns of the subset of factors discovered by our method but discarded by the frequentist method. These factors produce an annualized information ratio (IR) of 0.93 in the US and 1.10 globally (ex. US) over the full sample, with t -statistics above five. If we restrict analysis to the sample after that of Harvey et al. (2016b), the performance differential remains large and significant. These findings show strong external validity (post original publications, post Harvey et al. (2016b), different countries) and significant economic benefits of exploiting the joint information in all factor returns rather than simply applying a high cutoff for t -statistics.¹¹ We also show how the optimal risk-return profile has improved over time as factors have been discovered. In other words, the Sharpe ratio of the tangency portfolio has meaningfully increased over time as truly novel drivers of returns have been discovered. These findings can help inform asset pricing theory.

1 A Bayesian Model of Factor Replication

This section presents our Bayesian model for assessing factor replicability. We first draw out some basic implications of the Bayesian framework for interpreting evidence on individual factor alphas, then present a hierarchical structure for simultaneously modeling factors in a variety themes and across many countries.

1.1 Learning About Alpha: The Bayes Case

Posterior Alpha

We begin by considering an excess return factor f_t . A study of “anomalous” factor returns requires a risk benchmark, without which we cannot separate distinctive factor behavior from run of the mill risk compensation. We assume a CAPM benchmark due to its history

¹¹The out-of-sample performance across all significant factors under empirical Bayes is also highly significant as shown in Appendix Figure A2.

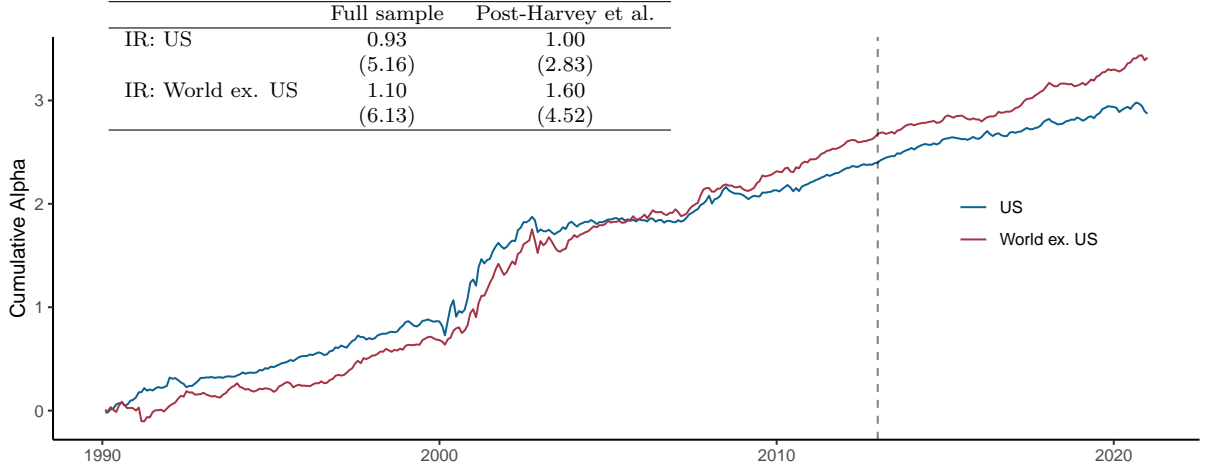


Figure 2. Out-of-sample performance of marginally significant factors

Note: The figure shows the cumulative CAPM alpha of an average of factors significant under our empirical Bayes framework, but not with the Benjamini-Yekutieli adjustment suggested by Harvey et al. (2016b). The significance cutoffs are re-estimated each year with the available data. Factors are eligible for inclusion after the sample period in the original paper, so all returns are out-of-sample. The table shows the information ratio (alpha divided by residual volatility) for the full sample (1990-2020) and the post-Harvey et al. (2016b) sample (2013-2020) with t -statistics in parentheses. The dashed line is at December 2012.

as a factor research benchmark for decades, and because it is not mechanically related to any of the factors that we attempt to replicate (in contrast to, say, the model of Fama and French, 1993b, which by construction explains size and value factors). A factor's net performance versus the excess market factor (r_t^m) is its α :

$$f_t = \alpha + \beta r_t^m + \varepsilon_t. \quad (2.1)$$

Our Bayesian prior is that the alpha is normally distributed with mean zero and variance τ^2 , or $\alpha \sim N(0, \tau^2)$. The mean of zero implies that CAPM holds on average, and τ governs potential deviations from CAPM. Intuitively, the higher the confidence in the prior, the lower is τ . The error term, $\varepsilon_t \sim N(0, \sigma^2)$, has volatility σ , is independent and identically distributed over time, and σ and β are observable.¹²

The risk-adjusted return, α , is estimated as the average market-adjusted factor return from T periods of data:

$$\hat{\alpha} = \frac{1}{T} \sum_t (f_t - \beta r_t^m) = \alpha + \frac{1}{T} \sum_t \varepsilon_t. \quad (2.2)$$

This observed ordinary least squares (OLS) estimate $\hat{\alpha}$ is distributed $N(\alpha, \sigma^2/T)$ given

¹²Here we seek to derive some simple expressions that illustrate the economic implications of Bayesian logic. In the empirical implementation, we use slightly richer model, as discussed further below. The empirical implementation normalizes factors so that σ is given at 10% for all factors, while β must be estimated, but this does not affect the economic points that we make in this section.

the true alpha, α . From Bayes' rule, we can compute the posterior distribution of the true alpha given the data evidence and prior. The posterior exhaustively describes the Bayesian's beliefs about alpha at a future time $t > T$ given the past experience, including the posterior expected factor performance,

$$E(\alpha|\hat{\alpha}) = E\left(f_t - \beta r_t^m \middle| \hat{\alpha}\right). \quad (2.3)$$

We derive the posterior alpha distribution via Bayes' rule (the derivation, which is standard, is shown in Appendix 5). The posterior alpha is normal with mean

$$E(\alpha|\hat{\alpha}) = \kappa \hat{\alpha} \quad (2.4)$$

where κ is a shrinkage factor given by

$$\kappa = \frac{\tau^2}{\tau^2 + \sigma^2/T} = \frac{1}{1 + \frac{\sigma^2}{\tau^2 T}} \in (0, 1) \quad (2.5)$$

and the posterior variance is

$$\text{Var}(\alpha|\hat{\alpha}) = \kappa \frac{\sigma^2}{T} = \frac{1}{\frac{1}{\sigma^2/T} + \frac{1}{\tau^2}}. \quad (2.6)$$

The first insight from this posterior is that a Bayesian predicts future returns will have smaller alpha (in absolute value) than the OLS estimate $\hat{\alpha}$, because the posterior mean ($\kappa \hat{\alpha}$) must lie between $\hat{\alpha}$ and the prior mean of zero. Said differently, a large observed alpha might be due to luck and, given the prior, we expect that at least part of this performance indeed is luck. The more data we have (higher T), the less shrinkage there is (i.e., κ closer to 1). Likewise, the stronger is the prior of zero alpha (i.e., lower τ), the heavier is the shrinkage. We can think of the prior τ in terms of the number of time periods of evidence that it corresponds to. That is, the posterior mean, $E(\alpha|\hat{\alpha})$, corresponds to first observing σ^2/τ^2 time periods with an average alpha of zero, followed by T time periods with a average alpha of $\hat{\alpha}$.

When evaluating out-of-sample evidence, a positive, but lower, alpha is sometimes interpreted as a sign of replication failure. But this is the expected outcome from the Bayesian perspective (i.e., based on the latest posterior), and can be fully consistent with a high degree of replicability. In fact, we show later that the comparatively low post-publication factor performance documented by McLean and Pontiff (2016) turns out to be consistent with the posterior a Bayesian would have formed given published results. Thus, post-publication results have tended to confirm the Bayesian's beliefs and as a

result the Bayesian posterior alpha estimate has been extraordinarily stable over time (see Section 3.2).

Alpha-hacking

Because out-of-sample alpha attenuation is not generally a sign of replication failure, we may want a more direct probe for non-replicability. We can build such a test into our Bayesian framework by embedding scope for “alpha-hacking,” or selectively reporting or manipulating data to artificially make the alpha seem larger. We represent this idea using the following distribution of factor returns in the in-sample time period $t = 1, \dots, T$:

$$f_t = \alpha + \beta r_t^m + \underbrace{\tilde{\varepsilon}_t + u}_{\varepsilon_t}. \quad (2.7)$$

Here, $\tilde{\varepsilon}_t \sim N(0, \sigma^2)$ captures usual return shocks and $u \sim N(\bar{\varepsilon}, \sigma_u^2)$ represents return inflation due to alpha-hacking. The total in-sample return shock ε_t is normally distributed, $N(\bar{\varepsilon}, \bar{\sigma}^2)$, where $\bar{\varepsilon} \geq 0$ is the alpha-hacking bias, and the variance $\bar{\sigma}^2 = \sigma^2 + \sigma_u^2 \geq \sigma^2$ is elevated due to the artificial noise created by alpha-hacking.¹³ Naturally, the false benefits of alpha-hacking disappear in out-of-sample data, or in other words $\varepsilon_t \sim N(0, \sigma^2)$ for $t > T$. The Bayesian accounts for alpha-hacking as follows:

Proposition 5 (Alpha-hacking) *The posterior alpha with alpha-hacking is given by*

$$E(\alpha|\hat{\alpha}) = -\kappa_0 + \kappa^{\text{hacking}}\hat{\alpha} \quad (2.8)$$

where $\kappa^{\text{hacking}} = \frac{1}{1 + \frac{\bar{\sigma}^2}{\tau^2 T}} \leq \kappa$ and $\kappa_0 = \kappa^{\text{hacking}}\bar{\varepsilon} \geq 0$. Further, $\kappa^{\text{hacking}} \rightarrow 0$ in the limit of “pure alpha-hacking,” $\tau \rightarrow 0$ or $\bar{\sigma} \rightarrow \infty$.

The Bayesian posterior alpha accounts for alpha-hacking in two ways. First, the estimated alpha is shrunk more heavily toward zero since the factor κ^{hacking} is now smaller. Second, the alpha is further discounted by the intercept term κ_0 due to the bias in the error terms.

We examine alpha-hacking empirically in Section 3.2 in light of Proposition 5. We consider a cross-sectional regression of factors’ out-of-sample (e.g., post-publication) alphas on their in-sample alphas, looking for the signatures of alpha-hacking in the form of a negative intercept term or a slope coefficient that is too small. In addition, Section 3.3 shows how to estimate the Bayesian model in a way that is less susceptible to the effects of alpha hacking and Appendix 5 presents additional theoretical results characterizing alpha-hacking.

¹³We note that this elevated variance cannot be detected by looking at the in-sample variance of residual returns since the alpha-hacking term u does not depend on time t .

1.2 Hierarchical Bayesian Model

Shared Alphas: The Case of Complete Pooling

We now embed a critical aspect of factor research into our Bayesian framework: Factors are often correlated and conceptually related to each other. For concreteness, we begin with a setting in which the researcher has access to “domestic” evidence in (2.1) as well as “global” evidence from an international factor, f_t^g , with known exposure β^g to the global market index r_t^g :

$$f_t^g = \alpha + \beta^g r_t^g + \varepsilon_t^g. \quad (2.9)$$

Here, we assume that the true alpha for this global factor is the same as the domestic alpha. In other words, we have complete “pooling” of information about alpha across the two samples. As an alternative interpretation, the researcher could have access to two related factors, say two different value factors in the same country, and assume that they have the same alpha because they capture the same investment principle.

The global shock, ε_t^g , is normally distributed $N(0, \sigma^2)$, and ε_t^g and ε_t are jointly normal with correlation ρ .¹⁴ The estimated alpha based on the global evidence is simply its market-adjusted return:

$$\hat{\alpha}^g = \frac{1}{T} \sum_t (f_t^g - \beta^g r_t^g). \quad (2.10)$$

To see the power of global evidence (or, more generally, the power of observing related strategies), we consider the posterior when observing both the domestic and global evidence.

Proposition 6 (The Power of Shared Evidence) *The posterior alpha given the domestic estimate, $\hat{\alpha}$, and the global estimate, $\hat{\alpha}^g$, is normally distributed with mean*

$$E(\alpha | \hat{\alpha}, \hat{\alpha}^g) = \kappa^g \left(\frac{1}{2} \hat{\alpha} + \frac{1}{2} \hat{\alpha}^g \right). \quad (2.11)$$

The global shrinkage parameter is

$$\kappa^g = \frac{1}{1 + \frac{\sigma^2}{\tau^2 T} \frac{1+\rho}{2}} \in [\kappa, 1] \quad (2.12)$$

which decreases with the correlation ρ , attaining the minimum value, $\kappa^g = \kappa$, when $\rho = 1$.

¹⁴The framework can be generalized to a situation where the global shocks have a different volatility and sample length. In this case, the Bayesian posterior puts more weight on the sample with lower volatility and longer sample.

The posterior variance is lower when observing both domestic and global evidence:

$$\text{Var}(\alpha|\hat{\alpha}) \geq \text{Var}(\alpha|\hat{\alpha}, \hat{\alpha}^g). \quad (2.13)$$

Naturally, the posterior depends on the average alpha observed domestically and globally. Furthermore, the combined alpha is shrunk toward the prior of zero. The shrinkage factor κ^g is smaller (heavier shrinkage) if the markets are more correlated because the global evidence provides less new information. With low correlation, the global evidence adds a lot of independent information, shrinkage is lighter, and the Bayesian becomes more confident in the data and less reliant on the prior. The proposition shows that, if a factor has been found to work both domestically and globally, then the Bayesian expects stronger out-of-sample performance than a factor that has only worked domestically (or has only been analyzed domestically).

Two important effects are at play here, and both are important for understanding the empirical evidence presented below: The domestic and global alphas are shrunk both toward *each other* and toward *zero*. For example, suppose that a factor worked domestically but not globally, say $\hat{\alpha} = 10\% > \hat{\alpha}^g = 0\%$. Then the overall evidence points to an alpha of $\frac{1}{2}\hat{\alpha} + \frac{1}{2}\hat{\alpha}^g = 5\%$, but shrinkage toward the prior results in a lower posterior, say, 2.5%. Hence, the Bayesian expects future factor returns in both regions of 2.5%. That shared alphas are shrunk together is a key feature of a *joint* model, and it generally leads to different conclusions than when factors are evaluated independently. Next we consider a perhaps more realistic model in which factors are only partially shrunk toward each other.

Hierarchical Alphas: The Case of Partial Pooling

We now consider several factors, numbered $i = 1, \dots, N$. Factor i has a true alpha given by

$$\alpha^i = c + w^i. \quad (2.14)$$

Here, c is the common component of all alphas, which has a prior distribution given by $N(0, \tau_c^2)$. Likewise, w^i is the idiosyncratic alpha component, which has a prior distribution given by $N(0, \tau_w^2)$, independent of c and across i . Said differently, we can imagine that nature first picks of the overall c from $N(0, \tau_c^2)$ and then picks the factor-specific α^i from $N(c, \tau_w^2)$.

This hierarchical model is a realistic compromise between assuming that all factor alphas are completely different (using equation (2.4) for each alpha separately) and assuming that they are all the same (using Proposition 6). Rather than assuming no pooling

or complete pooling, the hierarchical model allows factors to have a common component and an idiosyncratic component.

Suppose we observe factor returns of

$$f_t^i = \alpha^i + \beta^i r_t^m + \varepsilon_t^i \quad (2.15)$$

where ε_t^i are normally distributed with mean 0 and variance σ^2 and $\text{Cor}(\varepsilon_t^i, \varepsilon_t^j) = \rho \geq 0$ for all i, j .¹⁵ Computing the observed alpha estimates as above, $\hat{\alpha}^i = \frac{1}{T} \sum_t (f_t^i - \beta^i r_t^m)$, we derive the posterior in the following result.¹⁶

Proposition 7 (Hierarchical Alphas) *The posterior alpha of factor i given the evidence on all factors is normally distributed with mean*

$$E(\alpha^i | \hat{\alpha}^1, \dots, \hat{\alpha}^N) = \frac{1}{1 + \frac{\rho\sigma^2}{\tau_c^2 T} + \frac{\tau_w^2 + (1-\rho)\sigma^2/T}{\tau_c^2 N}} \hat{\alpha}^\cdot + \frac{1}{1 + \frac{(1-\rho)\sigma^2}{\tau_w^2 T}} \left(\hat{\alpha}^i - \frac{1}{1 + \frac{\tau_w^2 + (1-\rho)\sigma^2/T}{(\tau_c^2 + \rho\sigma^2/T)N}} \hat{\alpha}^\cdot \right), \quad (2.16)$$

where $\hat{\alpha}^\cdot = \frac{1}{N} \sum_j \hat{\alpha}^j$ is average alpha. When the number of factors N grows, the limit is

$$\lim_{N \rightarrow \infty} E(\alpha^i | \hat{\alpha}^1, \dots, \hat{\alpha}^N) = \frac{1}{1 + \frac{\rho\sigma^2}{\tau_c^2 T}} \hat{\alpha}^\cdot + \frac{1}{1 + \frac{(1-\rho)\sigma^2}{\tau_w^2 T}} (\hat{\alpha}^i - \hat{\alpha}^\cdot). \quad (2.17)$$

The posterior variance of factor i 's alpha using the information in all factor returns is lower than the posterior variance when looking at this factor in isolation:

$$\text{Var}(\alpha^i | \hat{\alpha}^1, \dots, \hat{\alpha}^N) < \text{Var}(\alpha^i | \hat{\alpha}^i). \quad (2.18)$$

The posterior variance is decreasing in N and, as $N \rightarrow \infty$, its limit is

$$\text{Var}(\alpha^i | \hat{\alpha}^1, \dots, \hat{\alpha}^N) \searrow \frac{\rho\sigma^2}{T} \frac{1}{1 + \frac{\rho\sigma^2}{\tau_c^2 T}} + \frac{(1-\rho)\sigma^2}{T} \frac{1}{1 + \frac{(1-\rho)\sigma^2}{\tau_w^2 T}}. \quad (2.19)$$

The main insight of this proposition is that having data on many factors is helpful for estimating the alpha of any of them. Intuitively, the posterior for any individual alpha depends on all of the other observed alphas because they are all informative about the

¹⁵Alternatively, we can write the error terms in a similar way to how we write the alphas in (2.14), namely $\varepsilon_t^i = \sqrt{\rho} \tilde{\varepsilon}_t + \sqrt{1-\rho} \tilde{\varepsilon}_t^i$, where $\tilde{\varepsilon}_t^i$ are idiosyncratic shocks that are independent across factors and of the common shock $\tilde{\varepsilon}_t$, with $\text{Var}(\tilde{\varepsilon}_t^i) = \text{Var}(\tilde{\varepsilon}_t) = \sigma^2$. We note that we require (the empirically realistic case) that $\rho \geq 0$ since we cannot have an arbitrarily large number of normal random variables with equal negative correlation (because the corresponding variance-covariance matrix would not be positive semi-definite for large enough N).

¹⁶The general hierarchical model is used extensively in the statistics literature, see, e.g., Gelman et al. (2013), but to our knowledge the results in Proposition 7 are not in the literature.

common alpha component. That is, the other observed alphas tell us whether alpha exists in general or, said another way, tell us if the CAPM appears to be violated in general. Further, the factor's own observed alpha tells us whether this specific factor appears to be especially good or bad. Using all of the factors jointly reduces posterior variance for all alphas. In summary, the joint model with hierarchical alphas has the dual benefits of identifying the common component in alphas and tightening confidence intervals by sharing information among factors.

To understand the proposition in more detail, consider first the (unrealistic) case in which all factor returns have independent shocks ($\rho = 0$). In this case, we essentially know the overall alpha when we see many uncorrelated factors. Indeed, the average observed alpha becomes a precise estimator of the overall alpha with more and more observed factors, $\hat{\alpha} \rightarrow c$. Since we essentially know the overall alpha in this limit, the first term in (2.17) becomes $1 \times \hat{\alpha}$ when $\rho = 0$ meaning that we don't need any shrinkage here. The second term is the outperformance of factor i above the average alpha, and this outperformance is shrunk toward our prior of zero. Indeed, the outperformance is multiplied by a number less than one, and this multiplier naturally decreases in the return volatility σ and decreases in our conviction in the prior (increases in τ_w).

The posterior variance is also intuitive in the case of $\rho = 0$. The posterior variance is clearly lower compared to only observing the performance of factor i itself:

$$\text{Var}(\alpha^i | \hat{\alpha}^1, \hat{\alpha}^2 \dots) = \frac{\sigma^2}{T} \frac{1}{1 + \frac{\sigma^2}{\tau_w^2 T}} < \frac{\sigma^2}{T} \frac{1}{1 + \frac{\sigma^2}{(\tau_c^2 + \tau_w^2) T}} = \text{Var}(\alpha^i | \hat{\alpha}^i) \quad (2.20)$$

based on (2.19) and (2.6). With partial pooling, the posterior variance decreases because the denominator on the left does not have τ_c^2 , reflecting that uncertainty about the general alpha has been eliminated by observing many factors.

In the realistic case where factor returns are correlated ($\rho > 0$), we see that both the average alpha $\hat{\alpha}$ and factor i 's outperformance $\hat{\alpha}^i - \hat{\alpha}$ are shrunk toward the prior of zero. This is because we cannot precisely estimate the overall alpha even with an infinite number of correlated factors—the correlated part never vanishes. Nevertheless, we still shrink the confidence interval, $\text{Var}(\alpha^i | \hat{\alpha}^1, \dots, \hat{\alpha}^N) \leq \text{Var}(\alpha^i | \hat{\alpha}^i)$, since more information is always better than less.

Multi-level Hierarchical Model

The model development to this point is simplified to draw out its intuition. Our empirical implementation is based on a more realistic (and slightly more complex) model that takes into account that factors naturally belong to different economic themes and to different regions.

In our global analysis, we have N different characteristic signals (e.g., book-to-market) across K regions, for a total of NK factors (e.g., US, developed, and emerging markets versions of book-to-market). Each of the N signals belongs to a smaller number of J theme clusters, where one cluster consists of various value factors, another consists of various momentum factors, and so on. One level of our hierarchical model allows for partially shared alphas among factors in the same theme cluster. Another level allows for commonality across regions among factors associated with the same underlying characteristic, capturing for example the connection between the book-to-market factor in different markets.

Mathematically, this means that an individual factor i has an alpha of

$$\alpha^i = \alpha^o + c^j + s^n + w^i. \quad (2.21)$$

Concretely, suppose factor $i \in \{1, \dots, NK\}$ is the book-to-market factor in the US region. Part of its alpha is driven by a component that is common to all factors, α^o , which we dogmatically fix at zero to be conservative. In addition, this factor i belongs to the value cluster $j \in \{1, \dots, J\}$, which contributes a cluster-specific alpha $c^j \sim N(0, \tau_c^2)$. Next, since factor i is based on book-to-market characteristic $n \in \{1, \dots, N\}$, it has an incremental signal-specific alpha of $s^n \sim N(0, \tau_s^2)$ that is shared across regions—e.g., it’s the common behavior among book-to-market factors regardless of geography. Finally, $w^i \sim N(0, \tau_w^2)$ is factor i ’s idiosyncratic alpha, namely the incremental alpha that is unique to the US version of book-to-market.

We write this model in vector form as¹⁷

$$\alpha = \alpha^o \mathbf{1}_{NK} + Mc + Zs + w \quad (2.22)$$

where $\alpha = (\alpha^1, \dots, \alpha^{NK})'$, $c = (c^1, \dots, c^J)'$, $s = (s^1, \dots, s^N)'$, $w = (w^1, \dots, w^{NK})'$, M is the $NK \times J$ matrix of cluster memberships, and Z is the $NK \times N$ matrix indicating the characteristic that factor i is based on. In particular, $M_{i,j} = 1$ if factor i is in cluster j and $M_{i,j} = 0$ otherwise. Likewise, $Z_{i,n} = 1$ if factor i is based on characteristic n and $Z_{i,n} = 0$ otherwise. This hierarchical model implies that the prior variance of alpha, denoted Ω , is¹⁸

$$\Omega = \text{Var}(\alpha) = MM'\tau_c^2 + ZZ'\tau_s^2 + I_{NK}\tau_w^2. \quad (2.23)$$

¹⁷The notation $\mathbf{1}_N$ refers to an $N \times 1$ vector of ones and I_N is the $N \times N$ identity matrix.

¹⁸Stated differently, each diagonal element of Ω is $\tau_c^2 + \tau_s^2 + \tau_w^2$. Further, if $i \neq k$, then the $(i, k)^{th}$ element of Ω is $\tau_c^2 + \tau_s^2$ if i and k are constructed from the same signal in the same cluster in different regions, it is τ_c^2 if i and k are constructed from different signals in the same cluster, and it is 0 if i and k are in different clusters.

In some cases, we analyze this model within a single region, $K = 1$ (for example, in our US-only analysis). In this case, there is no difference between signal-specific alphas and idiosyncratic alphas, so we collapse one level of the model by setting $\tau^s = 0$ and $s^n = 0$ for $n \in \{1, \dots, N\}$. In any case, the following result shows how to compute the posterior distribution of all alphas based on the prior uncertainty, Ω , and a general variance-covariance matrix of return shocks, $\Sigma = \text{Var}(\varepsilon)$. This result is at the heart of our empirical analysis.

Proposition 8 *In the multi-level hierarchical model, the posterior of the vector of true alphas is normally distributed with posterior mean*

$$E(\alpha|\hat{\alpha}) = (\Omega^{-1} + T\Sigma^{-1})^{-1} (\Omega^{-1}\mathbf{1}_{NK}\alpha_0 + T\Sigma^{-1}\hat{\alpha}) \quad (2.24)$$

and posterior variance

$$\text{Var}(\alpha|\hat{\alpha}) = (\Omega^{-1} + T\Sigma^{-1})^{-1}. \quad (2.25)$$

As noted above, we set the mean prior alpha to zero ($\alpha_0 = 0$) in our empirical implementation. This prior is based on economic theory and leads to a conservative shrinkage toward zero as seen in (2.24). We note that, in the data, the observed alphas are mostly positive, not centered around zero. However, these positive alphas are related to the way that factors are signed, namely according to the convention in the original paper, which almost always leads to a positive factor return in the original sample. However, if we view this signing convention as somewhat arbitrary, then a symmetry argument implies that a prior of zero is again natural. Said differently, factor means would be centered around zero if we changed signs arbitrarily, so our prior is agnostic about these signs.

1.3 Bayesian Multiple Testing and Empirical Bayes Estimation

Frequentist MT corrections embody a principle of conservatism that seeks to limit false discoveries, controlling the family-wise error rate (FWER) or the false discovery rate (FDR). Leading frequentist methods achieve this by widening confidence intervals and raising p -values, but do not alter the underlying point estimate.

Bayesian Multiple Testing

A large statistics literature describes how Bayesian modeling is effective for making reliable inferences in the face of multiple testing.¹⁹ Drawing on this literature, our hierarchical

¹⁹See Gelman et al. (2012); Berry and Hochberg (1999); Greenland and Robins (1991); Efron and Tibshirani (2002), among others. See Gelman (2016) for an intuitive, informal discussion of the topic.

model is a prime example of how Bayesian methods accomplish their MT correction based on two key model features.

First is the model prior, which imposes statistical conservatism in analogy to frequentist MT methods. It anchors the researcher’s beliefs to a sensible default (e.g., all alphas are zero) in case the data are insufficiently informative about the parameters of interest. Reduction of false discoveries is achieved first by shrinking estimates toward the prior. When there is no information in the data, the alpha point estimate is the prior mean and there are no false discoveries. As data evidence accumulates, posterior beliefs migrate away from the prior toward the OLS alpha estimate. In the process, discoveries begin to emerge, though they remain dampened relative to OLS. In the large data limit, Bayesian beliefs converge on OLS with no MT correction, which is justified because in the limit there are no false discoveries. In other words, the prior embodies a particularly flexible form conservatism—the Bayesian model decides how severe of an MT correction to make based on the informativeness of the data.

Second is the hierarchical structure that captures joint behavior of factors. Modeling factors jointly means that each alpha is shrunk toward its cluster mean (i.e., toward related factors), in addition to being shrunk toward the prior of zero. So, if we observe a cluster of factors in which most perform poorly, then this evidence reduces the posterior alpha even for the few factors with strong performance—another form of Bayesian MT correction. In addition to this Bayesian discovery control coming through shrinkage of the posterior mean alpha, the Bayesian confidence interval also plays an important role and changes as a function of the data. Indeed, having data on related factors leads to a contraction of the confidence intervals in our joint Bayesian model. So while alpha shrinkage often has the effect of reducing discoveries, the increased precision from joint estimation has the opposite effect of enhancing statistical power and thus increases discoveries.

In summary, a typical implementation of frequentist MT corrections estimates parameters independently for each factor, leaves these parameters unchanged, but inflates p -values to reduce the number of discoveries. In contrast, our hierarchical model leverages dependence in the data to efficiently learn about all alphas simultaneously. All data therefore helps to determine the center and width of each alpha’s confidence interval (Propositions 7 and 8). This leads to more precise estimates with “built-in” Bayesian MT correction.

Empirical Bayes Estimation

Given the central role of the prior, it might seem problematic that the severity of the Bayesian MT adjustment is at the discretion of the researcher. A powerful (and somewhat surprising) aspect of a hierarchical model is that the prior can be learned in part from the

data. This idea is formalized in the idea of “empirical Bayes (EB)” estimation, which has emerged as a major toolkit for navigating multiple tests in high-dimensional statistical settings (Efron, 2012).

The general approach to EB is to specify a multi-level hierarchical model, and then to use the dispersion of estimated effects within each level to learn about the prior parameters for that level. In our setting, the specific implementation of EB is dictated by Proposition 8. We first compute each factor’s abnormal return, $\hat{\alpha}$, as the intercept in a CAPM regression on the market excess return. Next, we set the overall alpha prior mean, α^o , to zero to enforce conservatism in our inferences.

From here, the benefits of EB kick in: The realized dispersion in alphas across factors helps to determine the appropriate level of conviction for the prior (that is, the appropriate values for τ_c^2 , τ_s^2 , and τ_w^2). For example, if we compute the average alpha for each cluster, \hat{c}^j (e.g., the average value alpha, the average momentum alpha, and so on), the cross-sectional variation in \hat{c}^j suggests that $\tau_c^2 \cong \frac{1}{J-1} \sum_{j=1}^J (\hat{c}^j - \hat{c})^2$. The same idea applies to τ_s^2 . Likewise, variation in observed alphas after accounting for hierarchical connections is informative about $\tau_w^2 \cong \frac{1}{NK-N-J} \sum_{i=1}^N (\hat{w}^i)^2$, where $\hat{w} = \hat{\alpha} - M\hat{c} - Z\hat{s}$.

The above variances illustrate the point that EB can help calibrate prior variances using the data itself. But those calculations are too crude, because they ignore sampling variation coming from the noise in returns, ε , which has covariance matrix Σ . Empirical Bayes estimates the prior variances by maximizing the prior likelihood function of the observed alphas, $\hat{\alpha} \sim N(0, \Omega(\tau_c, \tau_s, \tau_w) + \hat{\Sigma}/T)$, where the notation emphasizes that Ω depends on τ_c , τ_s , and τ_w according to (2.23). The likelihood function accounts for sampling variation through the a plug-in estimate of the covariance matrix of factor return shocks, $\hat{\Sigma}$.²⁰ We collect the resulting hyper-parameters in τ , that is, $\tau_c, \tau_s, \tau_w, \hat{\Sigma}$, and β^i .

Bayesian FDR and FWER

With the EB estimates (τ) on hand, we can compute the posterior distribution of the alphas from Proposition 8. From the posterior, we can in turn compute Bayesian versions of the FDR and FWER. Suppose that we consider a factor to be “discovered” if its z -score is greater than the critical value $\bar{z} = 1.96$:

$$\frac{E(\alpha^i | \hat{\alpha}^1, \dots, \hat{\alpha}^N, \tau)}{\sqrt{\text{Var}(\alpha^i | \hat{\alpha}^1, \dots, \hat{\alpha}^N, \tau)}} > \bar{z}. \quad (2.26)$$

²⁰We discuss the details of our EB estimation procedure in Appendix 5.3.

Equivalently, factor i is discovered if $p\text{-null}_i < 2.5\%$,²¹ where we use the posterior to compute

$$p\text{-null}_i = Pr(\alpha^i < 0 | \hat{\alpha}^1, \dots, \hat{\alpha}^N, \tau). \quad (2.27)$$

In words, $p\text{-null}_i$ is the posterior probability that the null hypothesis is true, which is the Bayesian version of a frequentist p -value. Said differently, it is the posterior probability of a “false discovery,” namely the probability that the true alpha is actually non-positive.

We can further compute the Bayesian FDR as:

$$\text{FDR}^{\text{Bayes}} = E \left(\frac{\sum_i 1_{\{i \text{ false discovery}\}}}{\sum_i 1_{\{i \text{ discovery}\}}} \middle| \hat{\alpha}^1, \dots, \hat{\alpha}^N, \tau \right) \quad (2.28)$$

where we condition on the data including at least one discovery (so the denominator is not zero), otherwise FDR is set to zero (see Benjamini and Hochberg, 1995).

The following proposition is a novel characterization of the Bayesian FDR, and shows that it is the posterior probability of a false discovery, averaged across all discoveries:

Proposition 9 (Bayesian FDR) *Conditional on the parameters of the prior distribution τ and data with at least one discovery, the Bayesian false discovery rate can be computed as:*

$$\text{FDR}^{\text{Bayes}} = \frac{1}{\#\text{discoveries}} \sum_{i \text{ discovery}} p\text{-null}_i. \quad (2.29)$$

and is bounded, $\text{FDR}^{\text{Bayes}} \leq 2.5\%$.

This result shows explicitly how the Bayesian framework controls the false discovery rate without the need for additional MT adjustments.²² The definition of a discovery ensures that at most 2.5% of the discoveries are false according to the Bayesian posterior, which is exactly the right distribution for assessing discoveries from the perspective of the Bayesian. Further, if many of the discovered factors are highly significant (as is the case in our data), then the Bayesian FDR is much lower than 2.5%.²³

We can also compute a Bayesian version of the family-wise error rate, which is the

²¹We use a critical value of 2.5% rather than 5% because the 1.96 cut-off corresponds to a 2-sided test, while false discoveries are only on one side in the Bayesian framework.

²²Efron (2007) includes related analysis but, to our knowledge, this particular result is new.

²³Proposition 9 formalizes the argument of Greenland and Robins (1991) that “from the empirical-Bayes or Bayesian perspective, multiple comparisons are not really a ‘problem.’ Rather, the multiplicity of comparisons provides an opportunity to improve our estimates through judicious use of any prior information (in the form of model assumptions) about the ensemble of parameters being estimated.”

probability of making one or more false discoveries in total:

$$\text{FWER}^{\text{Bayes}} = Pr \left(\sum_i 1_{\{i \text{ false discovery}\}} \geq 1 \mid \hat{\alpha}^1, \dots, \hat{\alpha}^N, \tau \right). \quad (2.30)$$

If we define a discovery as in (2.26) using the standard critical value $\bar{z} = 1.96$, then we do not necessarily control the family-wise error rate, $\text{FWER}^{\text{Bayes}}$, which is a harsh criterion that is concerned with the risk of a single false discovery without regard for the number of missed discoveries. $\text{FWER}^{\text{Bayes}}$ is a probability that can be computed from the posterior so it is straightforward to choose a critical value \bar{z} to ensure $\text{FWER}^{\text{Bayes}} \leq 5\%$ or any other level one prefers. The main point is that the Bayesian approach to replication lends itself to any inferential calculation the researcher desires because the posterior is a complete characterization of Bayesian beliefs about model parameters.

A Comparison of Frequentist and Bayesian False Discovery Control

We illustrate the benefits of Bayesian inference for our replication analysis via simulation. We assume a factor generating process based on the hierarchical model above and, for simplicity, consider a single region (as in our empirical US-only analysis), removing s^n and τ_s^2 from equations (2.21) and (2.23). We analyze discoveries as we vary the prior variances τ_c and τ_w . The remaining parameters are calibrated to our estimates for the US region in our empirical analysis below.

We simulate an economy with 130 factors in 13 different clusters of 10 factors each, observed monthly over 70 years. We assume that the mean alpha, α^o , is zero. We then draw a cluster alpha from $c^j \sim N(0, \tau_c^2)$ and a factor-specific alpha as $w^i \sim N(0, \tau_w^2)$. Based on these alphas, we generate realized returns by adding Gaussian noise.²⁴

We compute p -values separately using OLS with no adjustment or adjusting with the Benjamini-Yekutieli (BY) method. We also use EB to estimate the posterior alpha distribution, treating τ_c and τ_w as known in order to simplify simulations and focus on the Bayesian updating. For OLS and BY, a discovery occurs when the alpha estimate is positive and the two-sided p -value is below 5%. For EB, we consider it a discovery when the posterior probability that alpha is negative is less than 2.5%. For each pair of τ_c and τ_w , we draw 10,000 simulated samples, and report average discovery rates over all simulations.

Figure 3 reports alpha discoveries based on the OLS, BY, and EB approaches. For each method, we report the true FDR in the top panels (we know the truth since this is

²⁴The noise covariance matrix has a block structure calibrated to our data, with a correlation of 0.58 among factors in the same cluster and a correlation of 0.02 across clusters. The residual volatility for each factor is 10% per annum.

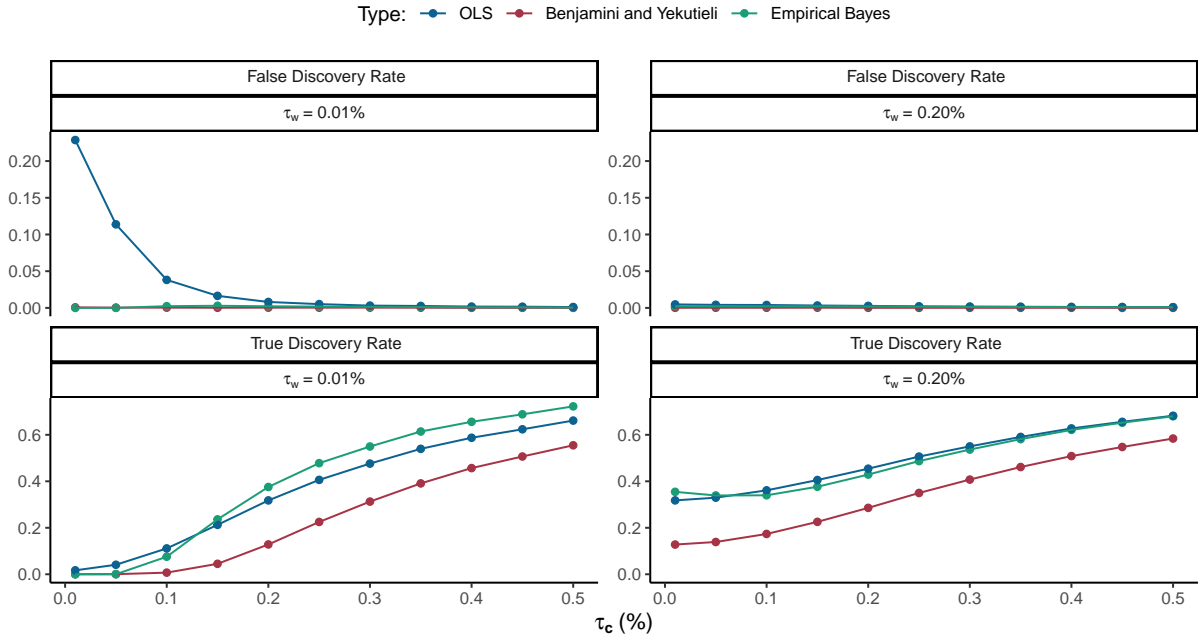


Figure 3. Simulation Comparison of False Discovery Rates

Note: The upper panels show the realized false discovery rate computed as the proportion of discovered factors for which the true alpha is negative, averaged over 10,000 simulations. The lower panels show the true discovery rate computed as the number of discoveries where the true alpha is positive divided by the total number of factors where the true alpha is positive. The left and right panels use low and high values of idiosyncratic variation in alphas (τ_w), respectively. The x -axis varies cluster alpha dispersion, τ_c .

a simulation) and the “true discovery rate”²⁵ in the bottom panels.

When idiosyncratic variation in true alphas is small (left panels with $\tau_w = 0.01\%$) and the variation in cluster alphas is also small (values of τ_c near zero on the horizontal axis), alphas are very small and true discoveries are unlikely. In this case, the OLS false discovery rate can be as high as 25% as seen in the upper left panel. However, both BY and EB successfully correct this problem and lower the FDR. The lower left panel shows that the BY correction pays a high price for its correction in terms of statistical power when τ_c is larger. In contrast, EB exhibits much better power to detect true positives while maintaining a similar false discovery control as BY. In fact, when there are more discoveries to be made in the data (as τ_c increases), EB becomes even more likely to identify true positives than OLS. This is due to the joint nature of the Bayesian model, whose estimates are especially precise compared to OLS due to EB’s ability to learn more efficiently from dependent data. This illustrates a point of Greenland and Robins (1991)

²⁵We define the true discovery rate to be the number of significantly positive alphas according to, respectively, OLS, BY, and EB divided by the number of truly positive alphas. Given our simulation structure, half of the alphas are expected to be positive in any simulation. Some of these will be small (i.e., economically insignificant) positives, so a testing procedure would require a high degree of statistical power to detect them. This is why the true discovery rate is below one even for high values of τ_c .

that “Unlike conventional multiple comparisons, empirical-Bayes and Bayes approaches will alter and can improve point estimates and can provide more powerful tests and more precise (narrower) interval estimators.” When the idiosyncratic variation is larger ($\tau_w = 0.20\%$), there are many more true discoveries to be made, so the false discovery rate tends to be low even for OLS with no correction. Yet in the lower right panel we continue to see the costly loss of statistical power suffered by the BY correction.

In summary, EB accomplishes a flexible MT adjustment by adapting to the data generating process. When discoveries are rare so that there is a comparatively high likelihood of false discovery, EB imposes heavy shrinkage and behaves similarly to the conservative BY correction. In this case, the benefit of conservatism costs little in terms of power exactly because true discoveries are rare. Yet when discoveries are more likely, EB behaves more like uncorrected OLS, giving it high power to detect discoveries and suffering little in terms of false discoveries because true positives abound.

The limitations of frequentist MT corrections are well studied in the statistics literature. Berry and Hochberg (1999) note that “these procedures are very conservative (especially in large families) and have been subjected to criticism for paying too much in terms of power for achieving (conservative) control of selection effects.” The reason is that, while inflating confidence intervals and p -values indeed reduces the discovery of false positives, it also reduces power to detect true positives.

Much of the discussion around MT adjustments in the finance literature fails to consider the loss of power associated with frequentist corrections. But, as Greenland and Hoffman (2019) point out, this tradeoff should be a first-order consideration for a researcher navigating multiple tests, and frequentist MT corrections tend to place an implicit cost on false positives that can be unreasonably large. Unlike some medical contexts for example, there is no obvious motivation for asymmetric treatment of false positives and missed positives in factor research. The finance researcher may be willing to accept the risk of a few false discoveries to avoid missing too many true discoveries. In statistics, this is sometimes discussed in terms of an (abstract) cost of Type I versus Type II errors,²⁶ but in finance we can make this cost concrete: We can look at the profit of trading on the discovered factors, where the cost of false discoveries is then the resulting extra risk and money lost (Section 3.3).

²⁶As Greenland and Robins (1991) point out, “Decision analysis requires, in addition to the likelihood function, a loss function, which indicates the cost of each action under the various possible values for the unknown parameter (benefits would be expressed as negative costs). Construction of a loss function requires one to quantify costs in terms of dollars, lives lost, or some other common scale.”

2 A New Public Data Set of Global Factors

We study a global dataset with 153 factors in 93 countries. In this section, we provide a brief overview of our data construction. We have posted the code along with extensive documentation detailing every implementation choice that we make for each factor.²⁷

Factors

The set of factors we study is based on the exhaustive list compiled by Hou et al. (2020b). They study 202 different characteristic signals from which they build 452 factor portfolios. The proliferation is due to treating 1, 6, and 12-month holding periods for a given characteristic as different factors, and due to their inclusion of both annual and quarterly updates of some accounting-based factors. In contrast, we focus on a 1-month holding period for all factors, and we only include the version that updates with the most recent accounting data (which could be either annual or quarterly). Lastly, we exclude a small number of factors for which data is not available globally. This gives us a set of 180 feasible global factors. For this set, we exclude factors based on industry or analyst data because they have comparatively short samples.²⁸ This leaves us with 138 factors. Finally, we add 15 factors studied in the literature that were not included in Hou et al. (2020b).

For each characteristic, we build the 1-month holding period factor return within each country as follows. First, in each country and month, we sort stocks into characteristic terciles (top/middle/bottom third) with breakpoints based on non-micro stocks in that country.²⁹ For each tercile, we compute its “capped value weight” return, meaning that we weight stocks by their market equity winsorized at the NYSE 80th percentile. This construction ensures that tiny stocks have tiny weights and any one mega stock does not dominate a portfolio, seeking to create tradable, yet balanced, portfolios.³⁰ The factor is then defined as the high-tercile return minus the low-tercile return, corresponding to the excess return of a long-short zero-net-investment strategy. The factor is long (short) the tercile identified by the original paper to have the highest (lowest) expected return.

We scale all factors such that their monthly idiosyncratic volatility is $10\%/\sqrt{12}$ (i.e.,

²⁷It is available at <https://jkpfactors.com/> and at <https://github.com/bkelly-lab/ReplicationCrisis>.

²⁸Global industry codes (GICS) are only available from 2000 and I/B/E/S data from the mid-1980’s (but coverage in early years is somewhat sparse).

²⁹Specifically, we start with all non-micro stocks in a country (i.e., larger than NYSE 20th percentile) and sort them into three groups of equal numbers of stocks based on the characteristic, say book-to-market. Then we distribute the micro-cap stocks into the three groups based on the same characteristic breakpoints. This process ensures that the non-micro stocks are distributed equally across portfolios, creating more tradable portfolios.

³⁰For robustness, Figure A1 of the appendix reports our replication results to using standard, uncapped value weights to construct factors.

10% annualized), which ensures cross-sectional stationarity and a priori that factors are similar in terms of their information ratio (i.e., appraisal ratio). Finally, we compute each factor's $\hat{\alpha}^i$ via an OLS regression on a constant and the corresponding region's market portfolio.

For a factor return to be non-missing, we require that it has at least 5 stocks in each of the long and short legs. We also require a minimum of 60 non-missing monthly observations for each country-specific factor for inclusion in our sample. When grouping countries into regions (US, developed ex. US, and emerging) we use the MSCI development classification as of January 7th 2021. When aggregating factors across countries, we use capitalization-weighted averages of the country-specific factors. For the developed and emerging market factors, we require that at least three countries have non-missing factor returns.

Clusters

We group factors into clusters using hierarchical agglomerative clustering (Murtagh and Legendre, 2014). We define the distance between factors as one minus their pairwise correlation and use the linkage criterion of Ward (1963). The correlation is computed based on CAPM-residual returns of US factors signed as in the original paper. Appendix Figure A15 shows the resulting dendrogram, which illustrates the hierarchical clusters identified by the algorithm. Based on the dendrogram, we choose 13 clusters that demonstrate a high degree of economic and statistical similarity. The cluster names indicate the types of characteristics that dominate each group: Accruals*, Debt Issuance*, Investment*, Leverage*, Low risk, Momentum, Profit Growth, Profitability, Quality, Seasonality, Size*, Short-Term Reversal*, and Value, where the star (*) indicates that these factors bet against the corresponding characteristic (e.g., accrual factors go long stocks with low accruals while shorting those with high accruals). Appendix Figure A16 shows that the average within-cluster pairwise correlation is above 0.5 for 9 out of 13 clusters, and Table AIII provides details on the cluster assignment, sign convention, and original publication source for each factor.

Data and Characteristics

Return data is from CRSP for the US (beginning in 1926) and from Compustat for all other countries (beginning in 1986 for most developed countries).³¹ All accounting data is from Compustat. For international data, all variables are measured in US dollars (based on exchange rates from Compustat) and excess returns are relative to the US treasury

³¹Appendix Table AV shows start date and other information for all countries included in our dataset.

bill rate. To alleviate the influence of data errors in the international data, we winsorize returns from Compustat at 0.1% and 99.9% each month.

We restrict our focus to common stocks that are identified by Compustat as the primary security of the underlying firm and assign stocks to countries based on the country of their exchange.³² In the US, we include delisting returns from CRSP. If a delisting return is missing and the delisting is for a performance-based reason, we set the delisting return to -30% following Shumway (1997). In the global data, delisting returns are not available, so all performance-based delistings are assigned a return of -30% .

We build characteristics in a consistent way, that sometimes deviates from the exact implementation used in the original reference. For example, for characteristics that use book equity, we always follow the method in Fama and French (1993b). Furthermore, we always use the most recent accounting data, whether annual or quarterly. Quarterly income and cash flow items are aggregated over the previous four quarters to avoid distortions from seasonal effects. We assume that accounting data is available four months after the fiscal period end. When creating valuation ratios, we always use the most recent price data following Asness and Frazzini (2013). Section 5.12 in the internet appendix contains a detailed documentation of our data set.

Empirical Bayes Estimation

We estimate the hyperparameters and the posterior alpha distributions of our Bayesian model via EB. Appendix 5.3 provides details on the EB methodology and the estimated parameters.

3 Empirical Assessment of Factor Replicability

We now report replication results for our global factor sample. We first present an internal validity analysis by studying US factors over the full sample. Then we analyze external validity in the global cross section and in the time series (post-publication factor returns).

3.1 Internal Validity

We report full sample performance of US factors in Figure 4. Each panel illustrates the CAPM alpha point estimate of each factor corresponding to the dot at the center of the vertical bars. Vertical bars represent the 95% confidence interval for each estimate. Bar colors differentiate between three types of factors. Blue shows factors that are significant

³²Compustat identifies primary securities in the US, Canada and rest of the world. This means that some firms can have up to three securities in our data set. In practice, the vast majority of firms (97%) only have one security in our sample at a given point in time.

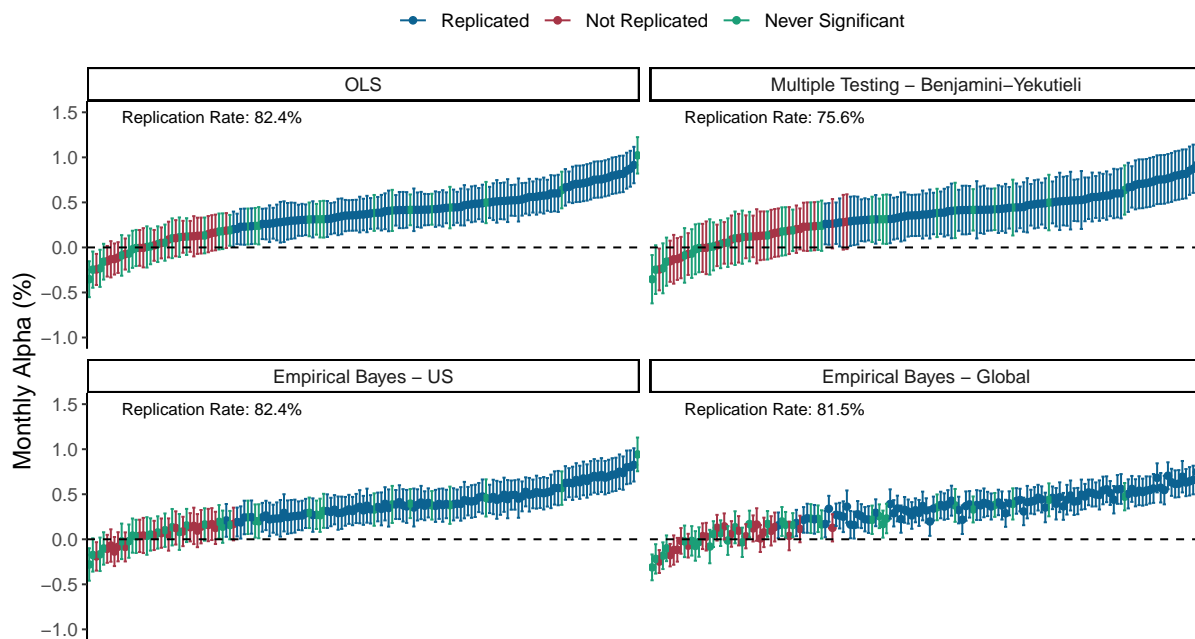


Figure 4. Alpha Distributions for US Factors

Note: The figure reports point estimates and confidence intervals for US factors. The upper left reports OLS estimates. The upper right uses the OLS point estimate but adjusts the confidence interval following the BY procedure. The lower left panel shows our EB posterior confidence intervals using only US data. The lower right continues to show EB results for US factors, but estimates the US factor posterior from global data rather than US-only data. Blue (red) confidence intervals correspond to factors that were significant in the original study in the literature and that we find significant (insignificant) based on the method in each panel. Green intervals correspond to factors that the original study find insignificant or do not evaluate in terms of average return significance. The order of factors is the same in all panels and is arranged from lowest OLS alpha to highest. Table AIV shows the factor names arranged in the same order.

in the original study and remain significant in our full sample. Red shows factors that are significant in the original study but are insignificant in our test. Green shows factors that are not significant in the original study, but are included in the sample of Hou et al. (2020b).

The four panels in Figure 4 differ in how the alphas and their confidence intervals are estimated. The upper left panel reports the simple OLS estimate of each alpha, $\hat{\alpha}_{ols}$, and the 95% confidence intervals based on unadjusted standard errors, $\hat{\alpha}_{ols} \pm 1.96 \times SE_{ols}$.³³ The factors are sorted by OLS $\hat{\alpha}$ estimate, and we use this ordering for the other three panels as well. We find that the OLS replication rate is 82.4%, computed as the number of blue factors (98) divided by the sum of red and blue factors (119). Based on OLS tests, factors are highly replicable.

The upper right panel repeats this analysis using the MT adjustment of Benjamini

³³We define SE_{ols} as the diagonal of the alpha covariance matrix $\hat{\Sigma}$, which we estimate according to Appendix 5.3.

and Yekutieli (2001) (denoted BY), which is advocated by Harvey et al. (2016b) and implemented by Hou et al. (2020b). This method leaves the OLS point estimate unchanged, but inflates the p -value. We illustrate this visually by widening the alpha confidence interval. Specifically, we find the BY-implied critical value³⁴ in our sample to be a t -statistic of 2.7, and we compute the corresponding confidence interval as $\hat{\alpha}_{\text{ols}} \pm 2.7 \times SE_{\text{ols}}$. We deem a factor as significant according to the BY method if this interval lies entirely above zero. Naturally, this widening of confidence intervals produces a lower replication rate of 75.6%. However, the BY correction does not materially change the OLS-based conclusion that factors appear highly replicable.

The lower left panel is based on our empirical Bayes estimates using the full sample of US factors. For each factor, we use Proposition 8 to compute its posterior mean, $E(\alpha_i | (\hat{\alpha}_j)_{j \text{ any US factor}})$, shown as the dot at the center of the confidence interval. These dots change relative to the OLS estimates, in contrast to BY and other frequentist MT methods that only change the size of the confidence intervals. We also compute the posterior volatility to produce Bayesian confidence intervals, $E(\alpha_i | (\hat{\alpha}_j)_{j \text{ any US factor}}) \pm 1.96 \times \sigma(\alpha_i | (\hat{\alpha}_j)_{j \text{ any US factor}})$. The replication rate based on Bayesian model estimates is 82.4%, larger than BY and, coincidentally, the same as the OLS replication rate. This replication rate has a built-in conservatism from the zero-alpha prior, and it further accounts for the multiplicity of factors because each factor’s posterior depends on *all* of the observed evidence in the US (not just own-factor performance).

The lower right panel again reports EB estimates for US factors, but now we allow the posterior to depend not just on US data, but on data from all over the world. That is, we compute the posterior mean and variance for each US factor conditional on the alpha estimates for all factors in all regions. The resulting replication rate is 81.5%, which is slightly lower than the EB replication rate using only US data. Some posterior means are reduced due to the fact that some factors have not performed as well outside the US, which affects posterior means for the US through the dependence among global alphas. For example, when the Bayesian model seeks to learn the true alpha of the “US change in book equity” factor, the Bayesian’s conviction regarding positive alpha is reduced by taking into account that the international version of this factor has underperformed the US version.³⁵

To further assess internal validity, we investigate the replication rate for US factors

³⁴We compute the BY-implied critical value as the average of the t -statistic of the factor that is just significant based on BY (the factor with the highest BY-adjusted p -value below 5%) and the t -statistic of the factor that is just insignificant (the factor with the lowest BY-adjusted p -value above 5%).

³⁵To provide a few more details on this example, the US factor based on annual change in book equity (be_gr1a) has a posterior volatility of 0.095% using only US data and 0.077% using global data, leading to a tighter confidence interval with the global data. However, the posterior mean is 0.22% using only US data and 0.13% using global data.

when those factors are constructed from subsamples based on stock size. One of the leading criticisms of factor research replicability is that results are driven by illiquid small stocks whose behavior in large part reflects market frictions and microstructure as opposed to just economic fundamentals or investor preferences. In particular, Hou et al. (2020b) argue that they find a low replication rate because they limit the influence of micro-caps. We find that factors demonstrate a high replication rate throughout the size distribution. Panel A of Figure 5 reports replication rates for US size categories shown in the five bars: mega stocks (largest 20% of stocks based on NYSE breakpoints), large stocks (market capitalization between the 80th and 50th percentile of NYSE stocks), small stocks (between the 50th and 20th percentile), micro stocks (between the 20th and 1st percentile), and nano stocks (market capitalization below the 1st percentile).

We see that the EB replication rates in mega and large stock samples are 77.3% and 79.8%, respectively. This is only marginally lower than the overall US sample replication rate of 82.4%, indicating that criticisms of factor replicability based on arguments around stock size or liquidity are largely groundless. For comparison, small, micro, and nano stocks deliver replication rates of 85.7%, 85.7% and 68.1%, respectively.

In Panel B of Figure 5, we report US factor replication rates by theme cluster. 11 out of 13 themes are replicable with a rate of 50% or better, with the exceptions being the low leverage and size themes. To understand these exceptions, we note that size factors are stronger in emerging markets (bottom panel of Figure A7) and among micro and nano stocks (bottom panels of Figure A8). The theoretical foundation of the size effect is a compensation for market illiquidity (Amihud and Mendelson, 1986) and market liquidity risk (Acharya and Pedersen, 2005). Theory predicts that the illiquidity (risk) premium should be the same order of magnitude as the differences in trading costs and these differences are simply much larger in emerging markets and among micro stocks.

Another reason why some factors and themes appear insignificant is that we are not accounting for other factors. Factors published after 1993 are routinely benchmarked to the Fama-French three-factor model (and, more recently, to the updated five-factor model). Some factors are insignificant in terms of raw return or CAPM alpha, but their alpha becomes significant after controlling for other factors. This indeed explains the lack of replicability for the low leverage theme. While CAPM alphas of low leverage factors are insignificant, we find that it is one of the best performing themes once we account for multiple factors (see Section 3.4 below).

3.2 External Validity

We find a high replication rate in our full-sample analysis, indicating that the large majority of factors are reproducible at least in-sample. We next study the external validity

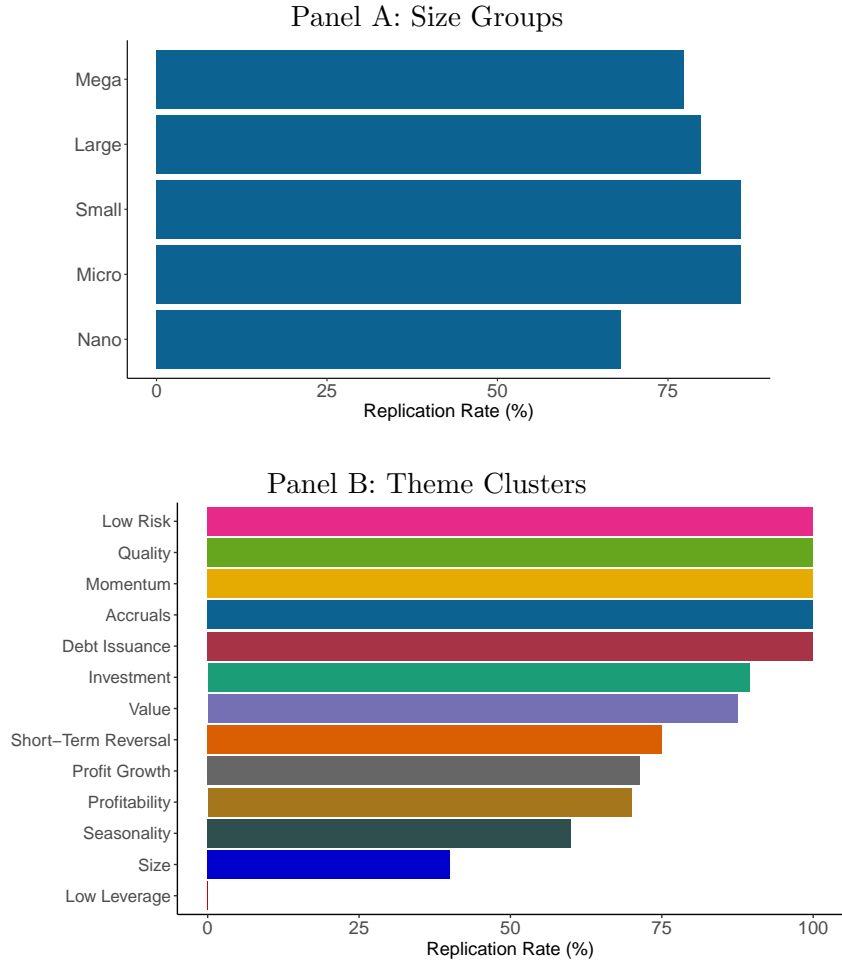


Figure 5. US Replication Rates By Size Group and Theme Cluster

Note: Panel A reports replication rates for US factors formed from subsamples defined by stocks’ market capitalization using our EB method. Panel B reports replication rates for US factors in each theme cluster.

of these results in international data and in post-publication US data.

Global Replication

Figure 6 shows corresponding replication rates around the world. We report replication rates from four testing approaches: (1) OLS with no adjustment; (2) OLS with Benjamini-Yekutieli MT adjustment; (3) the EB posterior conditioning only on factors within a region (“Empirical Bayes – Region”); and (4) EB conditioning on factors in all regions (“Empirical Bayes – All”). Even when using all global data to update the posterior of all factors, the reported Bayesian replication rate applies only to the factors within the stated region.

The first set of bars establishes a baseline by showing replication rates for the US sample, summarizing the results from Figure 4. The next two sets of bars correspond

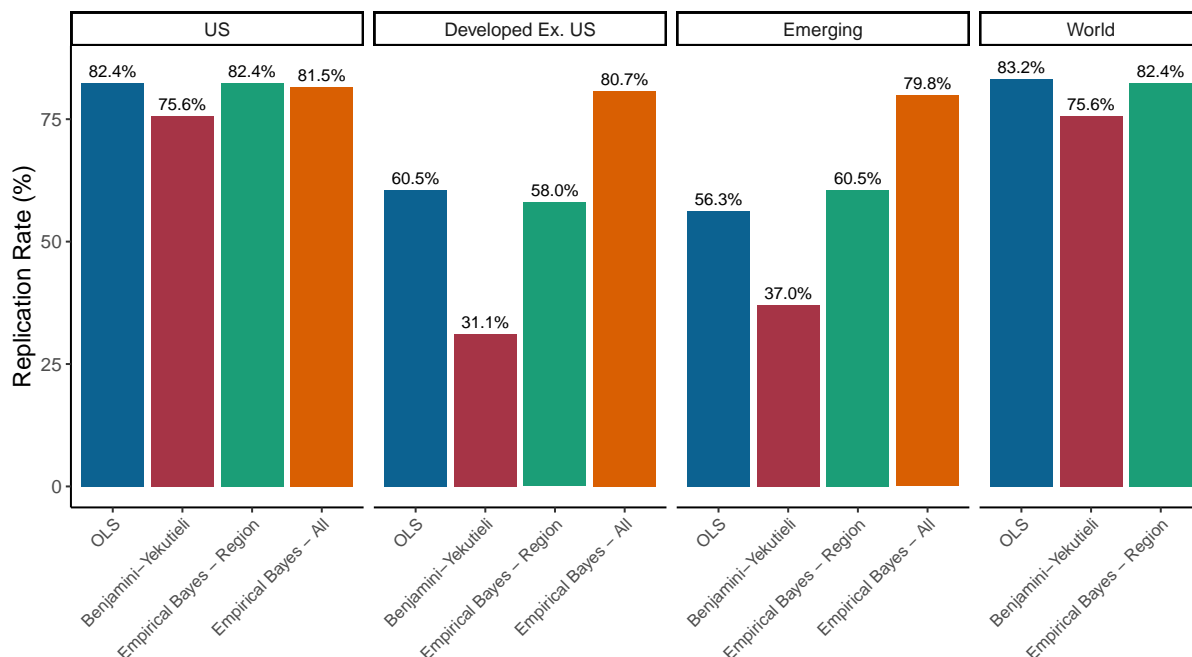


Figure 6. Replication Rates in Global Data

Note: We report replication rates for factors in three global regions (US, developed ex. US, and emerging) and for the world as a whole. A factor in a given region is the capitalization-weighted average factor for countries in that region. We report OLS replication rates with no adjustment and with Benjamini-Yekutieli multiple testing adjustment. We also report replication rates based on the empirical Bayes posterior. We consider two EB methods. In both methods, the replication rate refers only to factors within the region of interest, but the posterior is computed by conditioning either on data from that region alone (“Empirical Bayes – Region”) or on the full global sample (“Empirical Bayes – All”). We deem a factor successfully replicated if its 95% confidence interval excludes zero for a given method.

to the developed ex. US sample and the emerging markets sample, respectively.³⁶ Each region factor is a capitalization-weighted average of that factor among countries within a given region, and the replication rate describes the fraction of significant CAPM alphas for these regional factors.

OLS replication rates in developed and emerging markets are generally lower than in the US, and the frequentist Benjamini-Yekutieli correction has an especially large negative impact on replication rate. This is a case in which the Bayesian approach to MT is especially powerful. Even though the alphas of all regions are shrunk toward zero, the global information set helps EB achieve a high degree of precision, narrowing the posterior distribution around the shrunk point estimate. We can see this in increments. First, the EB replication rate using region-specific data (“Empirical Bayes – Region” in the figure) is just below the OLS replication rate but much higher than the Benjamini-Yekutieli rate.

³⁶The developed and emerging samples are defined by the MSCI development classification and include 23 and 27 countries, respectively. The remaining 43 countries in our sample that are classified as neither developed nor emerging by MSCI do not appear in our developed and emerging region portfolios, but they are included in the “world” versions of our factor portfolios.

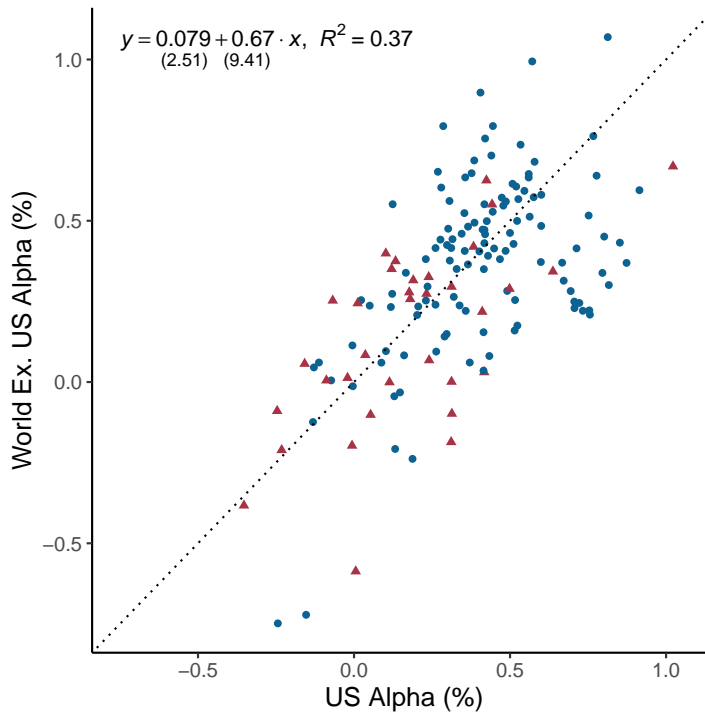


Figure 7. US Factor Alphas Versus World Ex. US

Note: The figure compares OLS alphas for US factors versus their international counterpart. Each world ex. US factor is a capitalization-weighted average of the factor in all other countries of our sample. Blue points correspond to factors that were significant in the original study in the literature, while red points are those for which the original paper did not find a significant effect (or did not study the factor in terms of average return significance). The dotted line is the 45° line. The figure also reports a regression of world ex. US alpha on US alpha.

When the posterior leverages global data (“Empirical Bayes – All” in the figure), the replication rate is higher still, reflecting the benefits of sharing information across regions, as recommended by the dependence among alphas in the hierarchical model.

Finally, we use the global model to compute, for each factor, the capitalization-weighted average alpha across all countries in our sample (“World” in the figure). Using data from around the world, we find a Bayesian replication rate of 82.4%.

Why do international OLS replication rates differ from the US? This is due primarily to the the fact that foreign markets have shorter time samples. Point estimates are similar in magnitude for the US and international data. Figure 7 shows the alpha of each US factor against the alpha of the corresponding factor for the world ex. US universe. The data cloud aligns closely with the 45° line, demonstrating the close similarity of alpha magnitudes in the two samples. But shorter international samples widen confidence intervals, and this is the primary driver of the drop in OLS replication rates outside the US.

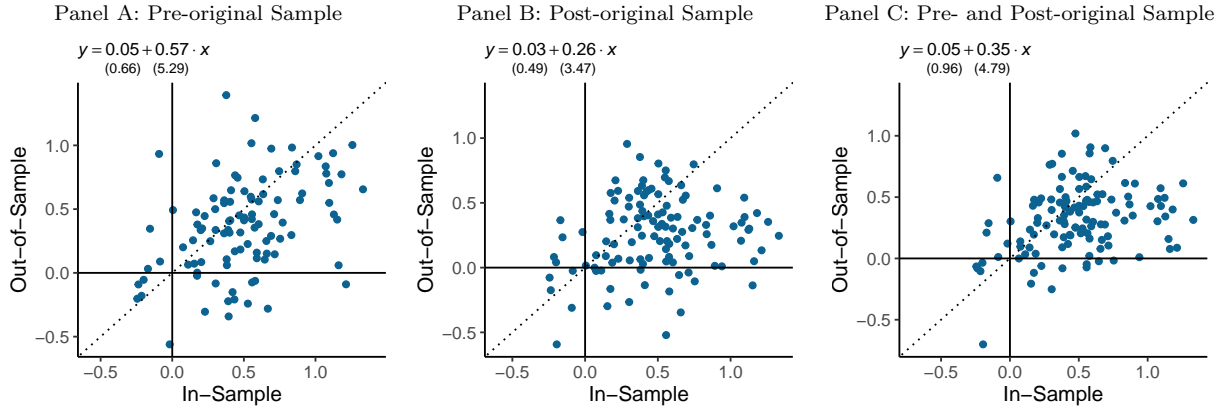


Figure 8. In-Sample versus Out-of-Sample Alphas for US Factors

Note: The figure plots OLS alphas for US factors during the in-sample period (i.e., the period studied in the original publication) versus out-of-sample alphas. In Panel A, out-of-sample is the time period before the in-sample period. In Panel B, out-of-sample is the time period before the in-sample period. In Panel C, out-of-sample includes both the time period before and after the in-sample period. We require at least five years of out-of-sample data for a factor to be included, amounting to 102, 115 and 119 factors in panel A, B and C. The figure also reports feasible GLS estimates of out-of-sample alphas on in-sample alphas. To implement feasible GLS, we assume that the error variance-covariance matrix is proportional to the full-sample CAPM residual variance-covariance matrix, $\hat{\Sigma}/T$, described in section 5.3. The dotted line is the 45° line.

Time Series Out-of-Sample Evidence

McLean and Pontiff (2016) document the intriguing fact that, following publication, factor performance tends to decay. They estimate an average post-publication decline of 58% in factor returns. In our data, the average in-sample alpha is 0.49% per month and the average out-of-sample alpha is 0.26% when looking post-original sample, implying a decline of 47%.

We gain further economic insight by looking at these findings cross-sectionally. Figure 8 makes a cross-sectional comparison of the in-sample and out-of-sample alphas of our US factors. The in-sample period is the sample studied in the original reference. The out-of-sample period in Panel A is the time period before the start of in-sample period, while in Panel B it is the period following the in-sample period. Panel C defines out-of-sample as the combined data from the periods before and after the originally studied sample. We find that 82.6% of the US factors that were significant in the original publication also have positive returns in the pre-original sample, 83.3% are positive in the post-original sample, and 87.4% are positive in the combined out-of-sample period. When we regress out-of-sample alphas on in-sample alphas using GLS, we find a slope coefficient of 0.57, 0.26, and 0.35 in Panels A, B, and C, respectively. The slopes are highly significant (ranging from $t = 3.5$ to $t = 5.3$) indicating that in-sample alphas contain something “real” rather than being the outcome of pure data mining, as factors that performed better in-sample also tend to perform better out-of-sample.

The significantly positive slope allows us to reject the hypothesis of “pure alpha-hacking,” which would imply a slope of zero, as seen in Proposition 5. Further, the regression intercept is positive, while alpha-hacking of the form studied in Proposition 5 would imply a negative intercept.

That the slope coefficient is positive and less than one is consistent with basic Bayesian logic of equation (2.4). As we emphasize in Section 1, a Bayesian would expect at least some attenuation in out-of-sample performance. This is because the published studies report the OLS estimate, while Bayesian beliefs shrink the OLS estimate toward the zero-alpha prior. More specifically, with no alpha hacking or arbitrage, the Bayesian expects a slope of approximately 0.9 using equation (2.5) and our EB hyperparameters (see appendix Table AI).³⁷ Hence, the slope coefficients in Figure 8 are too low relative to this Bayesian benchmark. In addition to the moderate slope, there is evidence that the dots in Figure 8 have a concave shape (as seen more clearly in appendix Figure A3). These results indicate that, while we can rule out pure alpha-hacking (or *p*-hacking), there is some evidence that the highest in-sample alphas may either be data-mined or arbitrated down.

From the Bayesian perspective, another interesting evaluation of time series external validity is to ask whether the new information contained in out-of-sample data moves the posterior alpha toward zero or not. Imagine a Bayesian observing the arrival of factor data in real time. As new data arrives, she updates her beliefs for all factors based on the information in the full cross section of factor data. In the top panel of Figure 9, we show how the Bayesian’s posterior of the average alpha would have evolved in real time.³⁸ We focus on all the World factors that are available since at least 1955 and significant in the original paper. Starting in 1960, we re-estimate the hierarchical model using the empirical Bayes estimator in December of each year. The plot shows the CAPM alpha and corresponding 95% confidence interval of an equal-weighted portfolio of the available factors. The posterior mean alpha becomes relatively stable from the mid 1980s, around 0.4% per month. And, as data evidence has accumulated over time, the confidence interval narrows by a third, from about 0.16% wide in 1960 to 0.10% in 2020.

To understand the posterior alpha, Figure 9 also shows the average OLS alpha as triangles and the bottom panel in Figure 9 reports the rolling 5-year average monthly alpha among all these factors. We see that the EB posterior is below the OLS estimate, which occurs because the Bayesian posterior is shrunk toward the zero prior. Naturally,

³⁷The slope is $\kappa = 1/(1 + \sigma^2/(T\tau^2)) = 0.9$, where $\sigma^2 = 10\%^2/12$, the average in-sample period length is $T = 420$ months, and $\tau^2 = \tau_c^2 + \tau_w^2 = (0.35\%)^2 + (0.21\%)^2 = (0.41\%)^2$.

³⁸Here we keep τ_c and τ_w fixed at their full-sample values of 0.37% and 0.23% to mimic the idea of given decision maker who starts with a given prior and updates this view based on new data, while keeping the prior fixed. Figure A4 shows that the figure is almost the same with rolling estimates of τ_c and τ_w , and Figure A5 shows that this consistency arises because the rolling estimates are relatively stable.

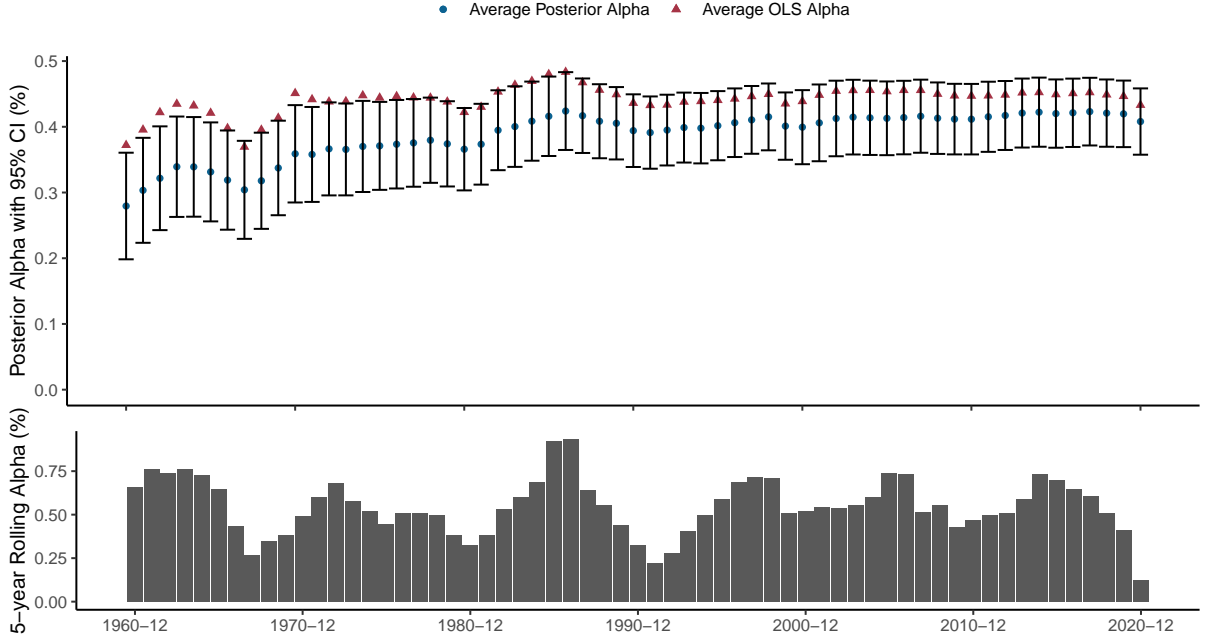


Figure 9. World Factor Alpha Posterior Distribution Over Time

Note: The top panel reports the CAPM alpha and 95% posterior confidence interval for an equal-weighted portfolio of World factors based on EB posteriors re-estimated in December each year. That is, each blue dot is $E(\frac{1}{N} \sum_i \alpha^i | \text{data until time } t)$ and the vertical lines are ± 2 times the posterior volatility. Triangles show average OLS alpha at each point in time, $\frac{1}{N} \sum_i \hat{\alpha}_{t_0, t}^i$, estimated using data through date t . The bottom panel reports the average monthly alpha for all factors in a rolling 5-year window. The results are based on factors found to be significant in the original paper with data available since 1955.

periods of good performance increase the posterior mean as well as the OLS estimate, and vice versa for poor performance. Over time, the OLS estimate moves nearer to the Bayesian posterior mean.

To further understand why the posterior alpha is relatively stable with a tightening confidence interval, consider the following simple example. Suppose a researcher has $T = 10$ years of data for factors with an OLS alpha estimate of $\hat{\alpha} = 10\%$ with standard error σ/\sqrt{T} . Further, assume their zero-alpha prior is equally as informative as their 10-year sample (i.e., $\tau = \sigma/\sqrt{T}$). Then the shrinkage factor is $\kappa = 1/2$ using equation (2.5). So, after observing the first ten years with $\hat{\alpha} = 10\%$, the Bayesian expects a future alpha of $E(\alpha|\hat{\alpha}) = 5\%$ (equation (2.4)). What happens if this Bayesian belief is confirmed by additional data, namely that the factor realizes an alpha of 5% over the next 10 years? In this case, the full-sample OLS of alpha is $\hat{\alpha} = 7.5\%$, but now the shrinkage factor becomes $\kappa = 2/3$ because the sample length doubles, $T = 20$. This results in a posterior alpha of $E(\alpha|\hat{\alpha}) = 7.5\% \cdot 2/3 = 5\%$. Naturally, when beliefs are confirmed by additional data, the posterior mean does not change. Nevertheless, we learn something from the additional data, because our conviction increases as the posterior variance is reduced. If $\sigma = 0.1$, the posterior volatility $\sqrt{\text{Var}(\alpha|\hat{\alpha})} = \sigma\sqrt{\frac{\kappa}{T}}$ goes from 2.2% with 10 years of data to 1.8%

with 20 years of data, and the confidence interval, $[E(\alpha|\hat{\alpha}) \pm 2\sqrt{\text{Var}(\alpha|\hat{\alpha})}]$, is reduced from [0.5%, 9.5%] to [1.3%, 8.7%].

3.3 Bayesian Multiple Testing

A great advantage of Bayesian methods for tackling challenges in multiple testing is that, from the posterior distribution, we can make explicit probability calculations for essentially any inferential question. We simulate from our EB posterior to investigate the false discovery and family-wise error rates among the set of global factors that were significant in the original study. We define a false discovery as a factor where we claim that the alpha is positive, but where the true alpha is negative.³⁹

First, based on Proposition 9, we calculate the Bayesian FDR in our sample as the average posterior probability of a false discovery, p -null, among all discoveries. We find that $\text{FDR}^{\text{Bayes}} = 0.1\%$, meaning that we expect roughly one discovery in 1000 to be a false positive given our Bayesian hierarchical model estimates. The posterior standard error for $\text{FDR}^{\text{Bayes}}$ is 0.3% with a confidence interval of [0,1%]. In other words, the model generates a highly conservative MT adjustment in the sense that once a factor is found to be significant, we can be relatively confident that the effect is genuine.

We can also use the posterior to make other inference calculations. We compute the FWER, which we define as the probability of at least one false discovery. We simulate 1,000,000 draws of the vector of alphas that were deemed to be discoveries from the EB posterior and compute

$$\text{FWER}^{\text{Bayes}} = \frac{1}{1,000,000} \sum_{s=1}^{1,000,000} 1_{\{n_s \geq 1\}} = 5.5\%$$

where n_s is the number of false discoveries in simulation s . In other words, the probability of at least one alpha having the wrong sign is 5.5%. The $\text{FWER}^{\text{Bayes}}$ is naturally much higher than the $\text{FDR}^{\text{Bayes}}$ given the extreme conservatism built into the FWER's definition of false discovery. Whether it is too high is subjective. A nice aspect of our approach is that a researcher can control the $\text{FWER}^{\text{Bayes}}$ as desired. For example, using a t -statistic threshold of 2.78 rather than 1.96 leads to $\text{FWER}^{\text{Bayes}} = 0.8\%$.

From the posterior, we can also compute the expected fraction of discovered factors that are "true," which is in general different than the replication rate. The replication rate is the fraction of factors having $E(\alpha_i|\text{data})/\sigma(\alpha_i|\text{data}) > 1.96$, while the expected fraction of true factors is $\frac{1}{n} \sum_i E(1_{\alpha_i > 0}|\text{data}) = \frac{1}{n} \sum_i Pr(\alpha_i > 0|\text{data})$. The replication rate gives

³⁹In particular, we define a discovery as a factor for which the posterior probability of the true alpha being negative is less than 2.5%. With this definition, we start with 153 world factors, then focus on the 119 factors that were significant in the original studies, and, out of these, 98 are considered discoveries.

a conservative take on the number of true factors—the expected fraction of true factors is typically higher than the replication rate. To understand this conservatism, consider an example in which all factors have a 90% posterior probability of being true. These would all individually be counted as “not replicated,” but they would contribute to a high expected fraction of true factors. Indeed, we estimate that the expected fraction of factors with truly positive alphas is 94% (with a posterior standard error of 1.3%), notably higher than our estimated replication rate.

Economic Benefits of More Powerful Tests

MT adjustments should ultimately be evaluated by whether they lead to better decisions. It is important to balance the relative costs of false positives versus false negatives, and the appropriate tradeoff depends on the context of the problem (Greenland and Hofman, 2019). We apply this general principle in our context by directly measuring costs in terms of investment performance.

Specifically, we can compute the difference in out-of-sample investment performance from investing using factors chosen with different methods. We compare two alternatives. One is the BY decision rule advocated by Harvey et al. (2016b), which is a frequentist MT method that successfully controls false discoveries relative to OLS, but in doing so sacrifices power (the ability to detect true positives). The second alternative is our EB method, whose false discovery control typically lies somewhere between BY and unadjusted OLS. EB uses the data sample itself to decide whether its discoveries should behave more similarly to BY or to unadjusted OLS.

For investors, the optimal decision rule is the one that leads to the best performance out-of-sample. For the most part, the set of discovered factors for BY and EB coincide. It is only in marginal cases where they disagree which, in our sample, occurs when EB makes a discovery that BY deems insignificant. Therefore, to evaluate MT approaches in economic terms, we track the out-of-sample performance of factors included by EB but excluded by BY. If the performance of these is negative on average, then the BY correction is warranted and preferred by the investor.

We find that the out-of-sample performance of factors discovered by EB but not BY is positive on average and highly significant. The alpha for these marginal cases is 0.35% per month among US factors ($t = 5.1$).⁴⁰ This estimate suggests that the BY decision rule is too conservative. An investor using the rule would fail to invest in factors that subsequently have a high out-of-sample return.

Another way to see that the BY decision rule is too conservative comes from the

⁴⁰For the developed ex. US sample, the monthly alpha for marginal cases is 0.24% per month ($t = 5.3$), and for the emerging sample it is 0.27% ($t = 3.7$), in favor of the EB decision rule. Appendix Table AII reports additional details for this analysis.

connection between Sharpe ratio and t -statistics: $t = \text{SR}\sqrt{T}$. If we have a factor with an annual Sharpe ratio of 0.5, an investor using the 1.96 cutoff would in expectation invest in the factor after 15 years. An investor using the 2.78 cutoff, would not start investing until observing the factor for 31 years.

Addressing Unobserved Factors, Publication Bias, and other Biases

A potential concern with our replication rate is that the set of factors that make it into the literature is a selected sample. In particular, researchers may have tried many different factors, some of which are observed in the literature, while others are unobserved because they never got published. Unobserved factors may have worse average performance if poor performance makes publication more difficult or less desirable. Alternatively, unobserved factors could have strong performance if people chose to trade on them in secret rather than publishing them. Either way, we next show how unobserved factors can be addressed in our framework.

The key insight is that the performance of factors across the universe of observed and unobserved factors is captured in our prior parameters τ_c, τ_w . Indeed, large values of these priors correspond to a large dispersion of alphas (that is, a lot of large alphas “out there”) while small values means that most true alphas are close to zero. Therefore, smaller τ ’s lead to a stronger shrinkage toward zero for our posterior alphas, leading to fewer factor “discoveries” and a lower replication rate. Figure 10 shows how our estimated replication rate depends on the most important prior parameter, τ_c , based on the τ_w that we estimated from the data.⁴¹

In Figure 10, we show how the replication rate varies with τ_c in precise quantitative terms. Note that while the replication rate indeed rises with τ_c , the differences are small in magnitude across a large range of τ_c values, demonstrating robustness of our conclusions about replicability.

This stable replication rate in Figure 10 also suggests that the replication rate among the observed factors would be similar even if we had observed the unobserved factors. The figure highlights several key values of τ_c : Both the value of τ_c that we estimated from the observed data (as explained in Appendix 5.3) and values that adjust for unobserved data in different ways.

We adjust τ_c for unobserved factors as follows. We simulate a data set that proxies for the full set of factors in the population (including those unobserved), and then estimate the τ ’s that match this sample. One set of simulations is constructed to match the baseline scenario of Harvey et al. (2016b) (Table 5.A, row 1), which estimates that researchers have tried $M = 1,297$ factors, of which 39.6% of have zero alpha and the rest have a Sharpe

⁴¹Figure A6 in the appendix shows that the results are robust to alternative values of τ_w .

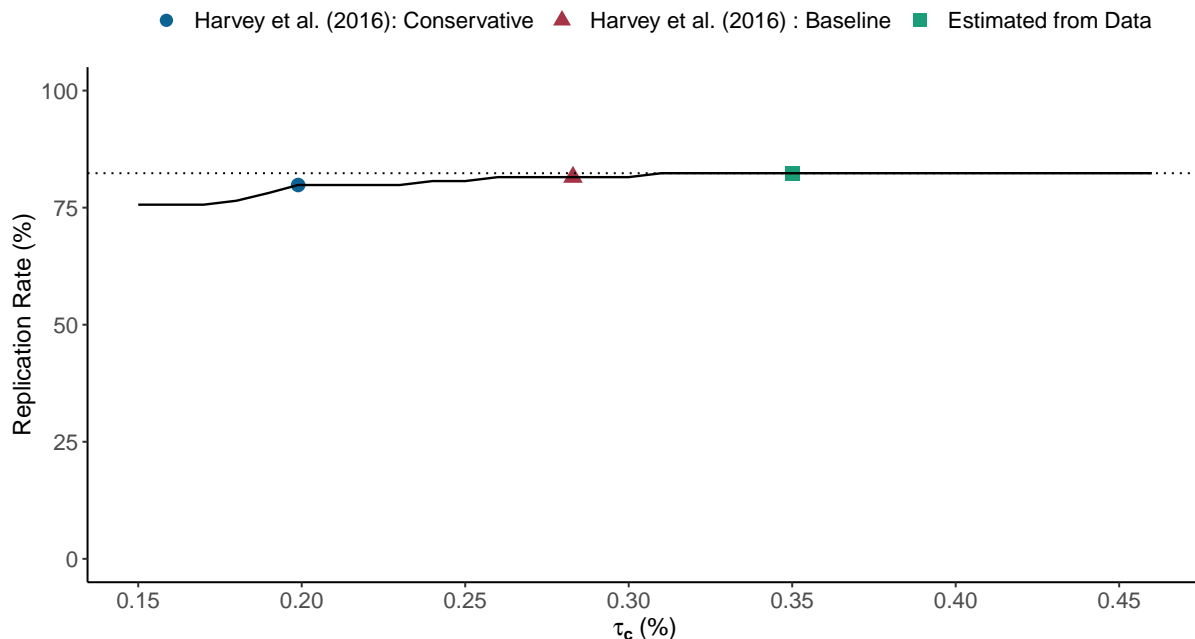


Figure 10. Replication Rate with Prior Estimated in Light of Unobserved Factors

Note: The figure shows how the replication rate in the US varies when changing the τ_c parameter. The τ_w parameter is fixed at the estimate value of 0.21%. The dotted line shows our replication rate of 82.4%. The green square, highlights the value estimated in the data $\tau_c = 0.35\%$. The red triangle and the blue circle highlights values that are found by estimating the empirical Bayes model according to assumptions about unobserved factors from Harvey et al. (2016b). The values are $\tau_c = 0.28\%$ in the baseline scenario and $\tau_c = 0.20\%$ in the conservative scenario. A description of this approach can be found in Appendix 5.7.

ratio of 0.44. We also consider the more conservative scenario of Harvey et al. (2016b) (Table 5.B, row 1), which implies that researchers have tried $M = 2,458$ factors, of which 68.3% have zero alpha. Appendix 5.7 has more details on these simulations. The result, as seen in Figure 10, is that values of τ_c that correspond to these scenarios from Harvey et al. (2016b) still lead to a conclusion of a high replication rate in our factor universe. The replication rate is 81.5%, and 79.8% for the prior hyperparameters implied by the baseline and conservative scenario respectively.

A closely related bias is that factors may suffer from alpha-hacking as discussed in Section 1.1 (Proposition 5), which makes realized in-sample factor returns too high. To account for this bias, we estimate the prior hyper-parameters using only out-of-sample data. These estimated values are $\tau_c = 0.27\%$ and $\tau_w = 0.22\%$. These hyper-parameters are similar to those implied by the baseline scenario of Harvey et al. (2016b) as seen in Figure 10. With these hyper-parameters, the replication rate is 81.5%.

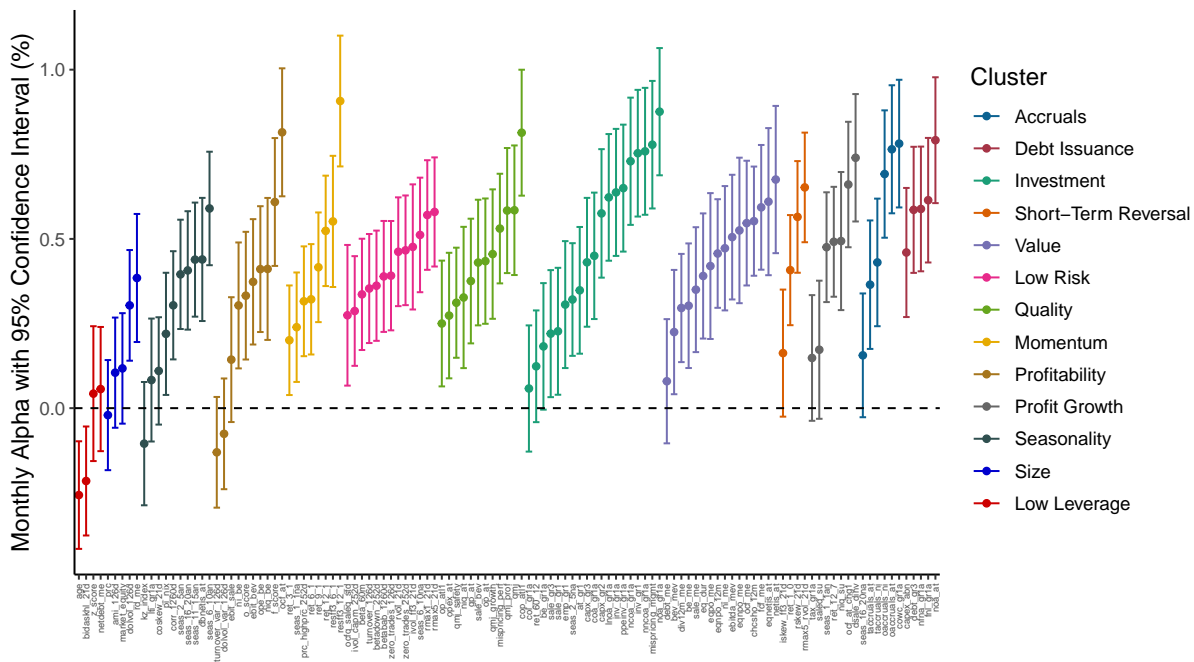


Figure 11. World Alpha Posterior By Factor and Cluster

Note: The figure reports the EB posterior 95% confidence interval for the true alpha of a world factor create as a capitalization weighted average of all country specific factors in our dataset. We only include factors that the original paper finds significant.

3.4 Economic Significance of Factors

Which factors (and which themes) are the most impactful anomalies in economic terms? We investigate this question by identifying which factors matter most from an investment performance standpoint.

Figure 11 shows the alpha confidence intervals for all world factors, sorted by the median posterior alpha within clusters. This illustration is similar to Figure 4, but now we focus on the world instead of the US factors, and here we sort factors into clusters. We also focus on factors that the original studies conclude are significant. We see that world factor alphas tend to be economically large, often above 0.3% per month, and tend to be highly significant, in most clusters. The exception is the low leverage cluster, where we also saw a low replication rate in preceding analyses.

By Region and By Size

We next consider which factors are most economically important across global regions and across stock size groups. In Panel A of Figure 12, we construct factors using only stocks in the five size subsamples presented earlier in Figure 5; namely mega, large, small, micro, and nano stock samples. For each sample, we calculate cluster-level alphas as

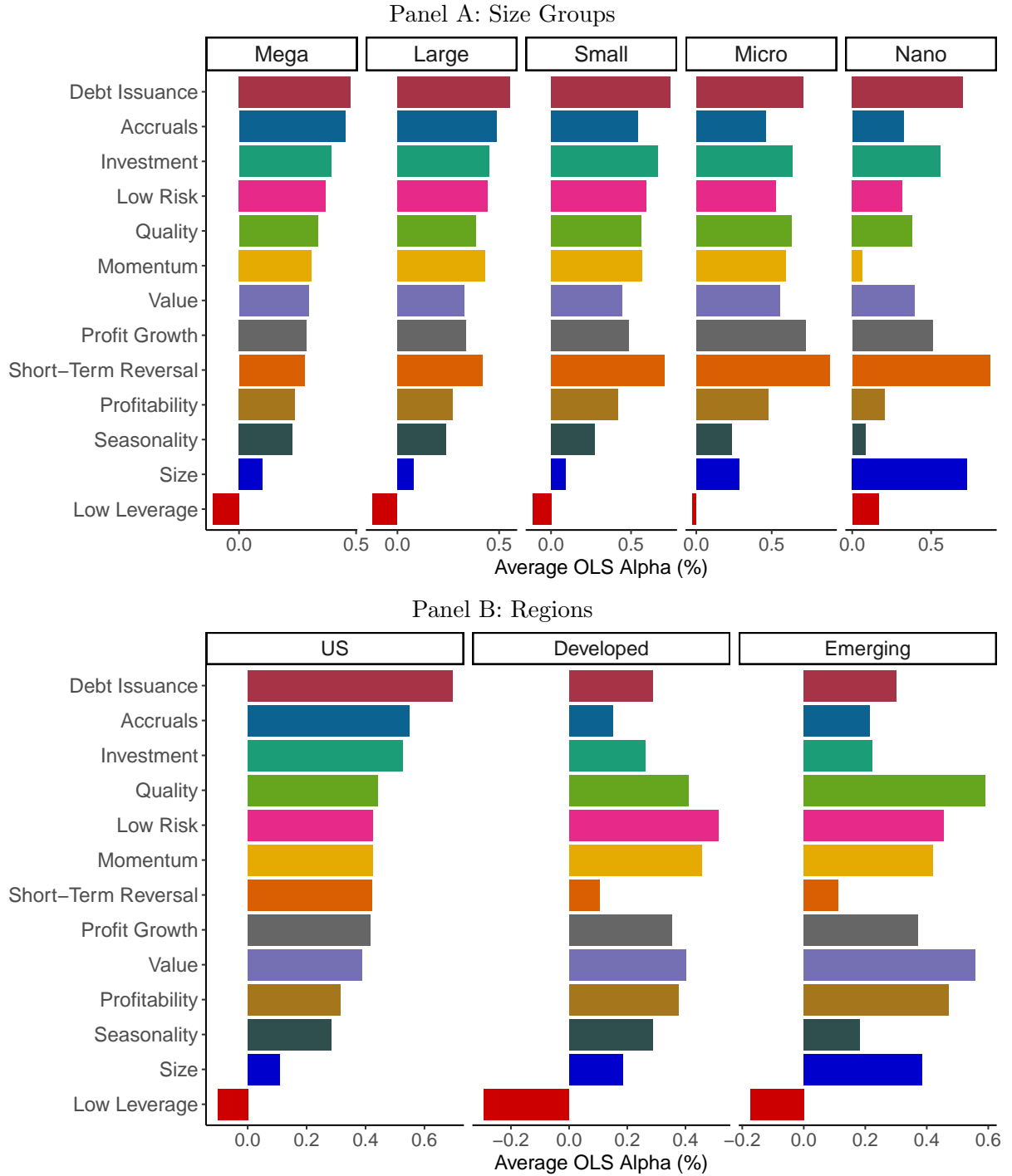


Figure 12. Alphas By Geographic Region and Stock Size Group

Note: The figure reports average cluster-level alphas for factors formed from subsamples defined by different stock market capitalization groups (Panel A) and regions (Panel B).

the equal-weighted average alpha of rank-weighted factors within the cluster.⁴² We see, perhaps surprisingly, that the ordering and magnitude of alphas is broadly similar across

⁴²Rank-weighting is similar to equal-weighting and used here to illustrate the performance of typical stocks in each size group. See equation (1) in Asness et al. (2013).

size groups. The Spearman rank correlation of alphas for mega caps versus micro caps is 73%. Only the nano stock sample, defined as stocks below the 1st percentile of the NYSE size distribution (which amounted to 458 out of 4356 stocks in the US at the end of 2020), exhibits notable deviation from the other groups. The Spearman rank correlation between alphas of mega caps and nano caps is 36%.

Panel B of Figure 12 shows cluster-level alphas across regions. Again, we find consistency in alphas across the globe, with the obvious standout being the size theme, which is much more important in emerging markets than in developed markets. US factor alphas share a 62% Spearman correlation with the developed ex. US sample, and a 43% correlation with the emerging markets sample.

Controlling for Other Themes

We have focused so far on whether factors (or clusters) possess significant positive alpha relative to the market. The limitation of studying factors in terms of CAPM alpha is that it does not control for duplicate behavior other than through the market factor. Economically important factors are those that have large impact on an investor's overall portfolio, and this requires understanding which clusters contribute alpha while controlling for all others.

To this end, we estimate cluster weights in a tangency portfolio that invests jointly in all cluster-level portfolios. We test the significance of the estimated weights using the method of Britten-Jones (1999). In addition to our 13 cluster-level factors, we also include the market portfolio as a way of benchmarking factors to the CAPM null. Lastly, we constrain all weights to be non-negative (because we have signed the factors to have positive expected returns according to the findings of the original studies).

Figure 13 reports the estimated tangency portfolio weights and their 90% bootstrap confidence intervals. When a factor has a significant weight in the tangency portfolio, it means that it matters for an investor, even controlling for all the other factors. We see that all but three clusters are significant in this sense. We also see that conclusions about cluster importance change when clusters are studied jointly. For example, value factors become stronger when controlling for other effects because of their hedging benefits relative to momentum, quality, and low leverage. More surprisingly, the low leverage cluster becomes one of the most heavily weighted clusters, in large part due to its ability to hedge value and low risk factors. The hedging performance of value and low leverage clusters is clearly discernible in Appendix table A16, which shows the average pairwise correlations among factors within and across clusters.⁴³ Appendix 5.9 provides further performance attribution of the tangency portfolio at the factor level.⁴⁴

⁴³Appendix Tables A9 and A10 show how tangency portfolio weights vary by region and by size group.

⁴⁴Figure A11 shows the performance of each cluster in combination with the market portfolio, figure

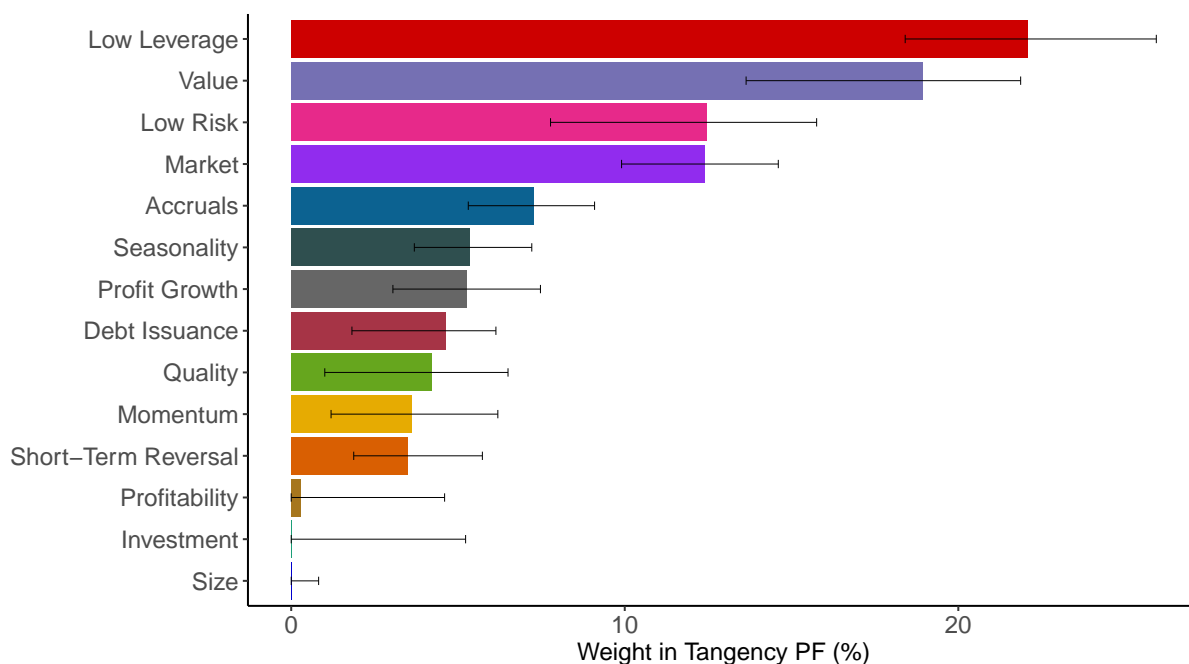


Figure 13. Tangency Portfolio Weights

Note: The return are from the US portfolios. We compute the cluster return as the equal weighted return of all factors with data available at a given point in time. We further add the US market return. We estimate the tangency weights following the method of Britten-Jones (1999) with a non-negativity constraint. The error bars are the 90% confidence intervals based on 10,000 bootstrap samples and the percentile method. The data starts in 1952 to ensure that all cluster have non-missing observations.

Evolution of Finance Factor Research

The number of published factors has increased over time as seen in the bottom panel of Figure 14. But, to what extent have these new factors continued to add new insights versus repackaging existing information?

To address this question, we consider how the optimal risk-return tradeoff has evolved over time as factors have been discovered. Specifically, Figure 14 computes the monthly Sharpe ratio of the ex-post tangency portfolio that only invests in factors discovered by a certain point in time.⁴⁵ The starting point (on the left) of the analysis is the 0.13 Sharpe ratio of the market portfolio in the US sample 1972-2020 when all factors are available. The ending point (on the right) is the 0.80 Sharpe ratio of the tangency portfolio that invests the optimal weights across all factors over the same US sample period.⁴⁶ In between, we see how the Sharpe ratio of the tangency portfolio has evolved

A12 shows how the optimal portfolio changes when one cluster is excluded and figure A14 shows the importance of each factor for the optimal portfolio.

⁴⁵We estimate tangency portfolio weights following the method of Pedersen (2021), which offers a sensible approach to mean-variance optimization for high dimensional data. Estimation details are provided in Appendix 5.9.

⁴⁶The high Sharpe ratio partly reflects the fact that we are doing an in-sample optimization. If we instead do a pseudo out-of-sample analysis via cross-validation, we find a monthly Sharpe ratio of 0.56.

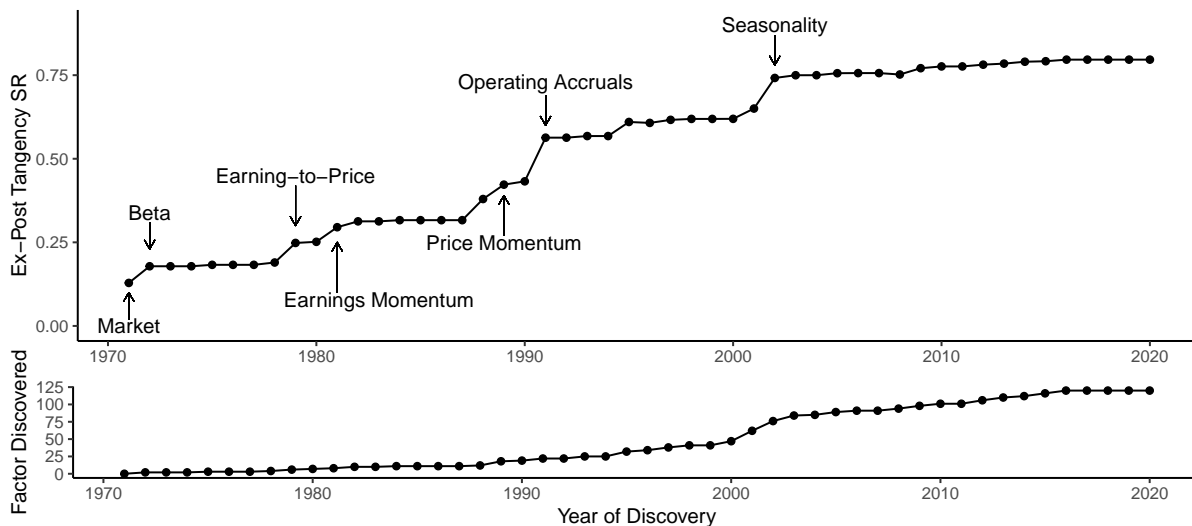


Figure 14. The Evolution of the Tangency Sharpe ratio

Note: The top panel shows the Sharpe ratio on the ex-post tangency portfolio. A factor is included in the tangency portfolio only after the end of the sample in which the factor was studied in the original publication (and we only include factors that were found to be significant in the original paper). We highlight selected factors that significantly improve the optimal portfolio, starting with the market portfolio. We use the longest available balanced US sample, 1972–2020 (that is, when all factors are available).

as factors have been discovered. The improvement is gradual over time, but we also see occasional large increases when researchers have discovered especially impactful factors (usually corresponding to new themes in our classification scheme). An example is the operating accruals factor proposed by Sloan (1996), which increased the tangency Sharpe ratio from 0.43 to 0.56. More recently, the seasonality factors of Heston and Sadka (2008) increase the Sharpe ratio from 0.65 to 0.74.

4 Conclusion: Finance Research Posterior

We introduce a hierarchical Bayesian model of alphas that emphasizes the joint behavior of factors and provides a more powerful multiple testing adjustment than common frequentist methods. Based on this framework, we re-visit the evidence on replicability in factor research and come to substantially different conclusions versus the prior literature. We find that US equity factors have a high degree of internal validity in the sense that over 80% of factors remain significant after modifications in factor construction that make all factors consistent, more implementable, while still capturing the original signal (Hamermesh, 2007) and after accounting for multiple testing concerns (Harvey et al., 2016b; Harvey, 2017).

We also provide new evidence demonstrating a high degree of external validity in factor research. In particular, we find highly similar qualitative and quantitative behavior in a large sample of 153 factors across 93 countries as we find in the US. We also show that, within the US, factors exhibit a high degree of consistency between their published in-sample results and out-of-sample data not considered in the original studies. We show that some out-of-sample factor decay is to be expected in light of Bayesian posteriors based on publication evidence. Therefore, the new evidence from post-publication data largely confirms the Bayesian's beliefs, which has led to relatively stable Bayesian alpha estimates over time.

In addition to providing a powerful tool for replication, our Bayesian framework has several additional applications. For example, the model can be used to correctly interpret out-of-sample evidence, look for evidence of alpha-hacking, compute the expected number of false discoveries and other relevant statistics based on the posterior, analyze portfolio choice taking into account both estimation uncertainty and return volatility, and evaluate asset pricing models.

Finally, the code, data, and meticulous documentation for our analysis are available online. Our large global factor data set and the underlying stock-level characteristics are easily accessible to researchers by using our publicly available code and its direct link to WRDS. Our database will be updated regularly with new data releases and code improvements. We hope that our methodology and data will help promote credible finance research.

5 Appendix

5.1 Additional Results on Alpha Hacking

We consider the situation where the researcher has in-sample data from time 1 to time T and an out-of-sample (oos) period from time $T + 1$ to $T + T^{oos}$. The researcher may have used alpha-hacking during the in-sample period, but this does not affect the out-of-sample period. The researcher is interested in the posterior alpha based on the total evidence, in-sample and out-of-sample, which is useful for predicting factor performance in a future time period (that is, a time period that is out-of-sample relative to the existing out-of-sample period).

Proposition 10 (Out-of-sample alpha) *The posterior alpha based on an in-sample data from time 1 to T with alpha-hacking, and an out-of-sample period from $T + 1$ to $T + T^{oos}$ is given by*

$$E(\alpha|\hat{\alpha}, \hat{\alpha}^{oos}) = \kappa^{oos} (w(\hat{\alpha} - \bar{\varepsilon}) + (1 - w)\alpha^{oos}) \quad (2.31)$$

where $w = \frac{\sigma^2/T^{oos}}{\bar{\sigma}^2/T + \sigma^2/T^{oos}} \in (0, 1)$ is the relative weight on the in-sample period relative to the out-of-sample period, and $\kappa^{oos} = \frac{1}{1 + 1/(\tau^2([\bar{\sigma}^2/T]^{-1} + [\sigma^2/T^{oos}]^{-1}))}$ is a shrinkage parameter.

We see that, the more alpha hacking the researcher has done (higher $\bar{\sigma}$), the less weight we put on the in-sample period relative to the out-of-sample period. Further, the in-sample period has the non-proportional discounting due to alpha hacking ($\bar{\varepsilon}$), which we don't have for out-of-sample evidence.

So this result formalizes the idea that an in-sample backtest plus live performance is *not* the same as a longer backtest. For example, 10 years of backtest plus 10 years of live performance is more meaningful than 20 years of backtest with no live performance. The difference is that the oos-performance is free from alpha-hacking.

5.2 Proofs and Lemmas

The proofs make repeated use of the following well-known property of multivariate Normally distributed random variable. If x and y are multivariate Normal:

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim N \left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} \Sigma_{xx} & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_{yy} \end{bmatrix} \right) \quad (2.32)$$

then the conditional distribution of x given y has the following Normal distribution:

$$x|y \sim N \left(\mu_x + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y), \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} \right) \quad (2.33)$$

The proofs also make use of the following two lemmas.

Lemma 1 For random variables x, y, z , it holds that $E(\text{Var}(x|y, z)) \leq E(\text{Var}(x|y))$ and, if the random variables are jointly normal, then $\text{Var}(x|y, z) \leq \text{Var}(x|y)$.

Lemma 2 Let A be an $N \times N$ matrix for which all diagonal elements equal a and all off-diagonal elements equal b , where $a \neq b$ and $a + b(N - 1) \neq 0$. Then the inverse A^{-1} exists and is of the same form:

$$A = \begin{bmatrix} a & & b \\ & \ddots & \\ b & & a \end{bmatrix} \quad A^{-1} = \begin{bmatrix} c & & d \\ & \ddots & \\ d & & c \end{bmatrix} \quad (2.34)$$

where $c = \frac{a+b(N-2)}{(a-b)(a+b(N-1))}$ and $d = \frac{-b}{(a-b)(a+b(N-1))}$.

Proof of Lemma 1. Using the definition of conditional variance, we have

$$E(\text{Var}(x|y, z)) = E(E(x^2|y, z)) - E([E(x|y, z)]^2) = E(x^2) - E([E(x|y, z)]^2)$$

Hence, using Jensen's inequality, we have

$$\begin{aligned} E(\text{Var}(x|y)) - E(\text{Var}(x|y, z)) &= E([E(x|y)]^2) - E([E(x|y, z)]^2) \\ &= E([E(x|y, z)]^2) - E([E(E(x|y, z)|y)]^2) \\ &\geq E([E(x|y, z)]^2) - E(E([E(x|y, z)]^2|y)) = 0 \end{aligned}$$

The result for normal distributions follows from the fact that normal conditional variances are non-stochastic, i.e., $\text{Var}(x|y) = E(\text{Var}(x|y))$. In this case, we can also characterize the extra drop in variance due to conditioning on z using its orthogonal component ε from the regression $z = a + by + \varepsilon$, using similar notation as (2.32):

$$\begin{aligned} \text{Var}(x|y, z) = \text{Var}(x|y, \varepsilon) &= \Sigma_{x,x} - \Sigma_{x,(y,\varepsilon)} \Sigma_{(y,\varepsilon),(y,\varepsilon)}^{-1} \Sigma_{(y,\varepsilon),x} \\ &= \Sigma_{x,x} - \Sigma_{x,y} \Sigma_{y,y}^{-1} \Sigma_{y,x} - \Sigma_{x,\varepsilon} \Sigma_{\varepsilon,\varepsilon}^{-1} \Sigma_{\varepsilon,x} = \text{Var}(x|y) - \Sigma_{x,\varepsilon} \Sigma_{\varepsilon,\varepsilon}^{-1} \Sigma_{\varepsilon,x} \end{aligned}$$

■

Proof of Lemma 2. The proof follows from inspection: The product of A and its proposed inverse clearly has the same form as A with diagonal elements

$$ac + bd(I - 1) = \frac{a(a + b(N - 2)) - b^2(N - 1)}{(a - b)(a + b(N - 1))} = \frac{a^2 + ab(N - 1) - ab - b^2(N - 1)}{(a - b)(a + b(N - 1))} = 1$$

and off-diagonal elements

$$ad + bc + bd(N - 2) = \frac{-ab + b(a + b(N - 2)) - b^2(N - 2)}{(a - b)^2(a + b(N - 1))^2} = 0$$

In other words, AA^{-1} equals the identity, proving the result. \blacksquare

Proof of Equations (2.4)–(2.6). The posterior distribution of the true alpha given the observed factor return is computed using (2.33). The conditional mean is

$$E(\alpha|\hat{\alpha}) = 0 + \frac{\text{Cov}(\alpha, \hat{\alpha})}{\text{Var}(\hat{\alpha})}(\hat{\alpha} - 0) = \frac{\tau^2}{\tau^2 + \sigma^2/T}\hat{\alpha} = \kappa\hat{\alpha}$$

where κ is given by (2.5) and the posterior variance is

$$\text{Var}(\alpha|\hat{\alpha}) = \text{Var}(\alpha) - \frac{(\text{Cov}(\alpha, \hat{\alpha}))^2}{\text{Var}(\hat{\alpha})} = \tau^2 - \tau^2 \frac{\tau^2}{\tau^2 + \sigma^2/T} = \frac{\tau^2\sigma^2/T}{\tau^2 + \sigma^2/T} = \kappa \frac{\sigma^2}{T}$$

\blacksquare

Proof of Proposition 5. The posterior alpha with alpha-hacking is given via (2.33) as

$$E(\alpha|\hat{\alpha}) = 0 + \frac{\text{Cov}(\alpha, \hat{\alpha})}{\text{Var}(\hat{\alpha})}(\hat{\alpha} - E(\hat{\alpha})) = \frac{\tau^2}{\tau^2 + \bar{\sigma}^2/T}(\hat{\alpha} - \bar{\varepsilon}) = -\kappa_0 + \kappa^{\text{hacking}}\hat{\alpha}$$

where $\kappa^{\text{hacking}} = \frac{1}{1 + \frac{\bar{\sigma}^2}{\tau^2 T}}$, $\kappa_0 = \kappa^{\text{hacking}}\bar{\varepsilon} \geq 0$, and $\kappa^{\text{hacking}} \leq \kappa$ because $\bar{\sigma} \geq \sigma$. \blacksquare

Proof of Proposition 6. The posterior mean given $\hat{\alpha}$ and $\hat{\alpha}^g$ is computed via (2.33) as

$$\begin{aligned} E(\alpha|\hat{\alpha}, \hat{\alpha}^g) &= \begin{bmatrix} \tau^2 & \tau^2 \end{bmatrix} \begin{bmatrix} \tau^2 + \sigma_T^2 & \tau^2 + \rho\sigma_T^2 \\ \tau^2 + \rho\sigma_T^2 & \tau^2 + \sigma_T^2 \end{bmatrix}^{-1} \begin{bmatrix} \hat{\alpha} \\ \hat{\alpha}^g \end{bmatrix} \\ &= \frac{1}{\det \begin{bmatrix} \tau^2 & \tau^2 \\ \tau^2 + \rho\sigma_T^2 & \tau^2 + \sigma_T^2 \end{bmatrix}} \begin{bmatrix} \tau^2 + \sigma_T^2 & -(\tau^2 + \rho\sigma_T^2) \\ -(\tau^2 + \rho\sigma_T^2) & \tau^2 + \sigma_T^2 \end{bmatrix} \begin{bmatrix} \hat{\alpha} \\ \hat{\alpha}^g \end{bmatrix} \\ &= \frac{\tau^2(1 - \rho)\sigma_T^2}{\det} (\hat{\alpha} + \hat{\alpha}^g) \\ &= \frac{\tau^2(1 - \rho)}{\sigma_T^2(1 - \rho)(1 + \rho) + 2\tau^2(1 - \rho)} (\hat{\alpha} + \hat{\alpha}^g) \\ &= \kappa^g \left(\frac{1}{2}\hat{\alpha} + \frac{1}{2}\hat{\alpha}^g \right) \end{aligned}$$

using the notation $\sigma_T^2 = \sigma^2/T$ and

$$\det = (\tau^2 + \sigma_T^2)^2 - (\tau^2 + \rho\sigma_T^2)^2 = \sigma_T^2[\sigma_T^2(1 - \rho^2) + 2\tau^2(1 - \rho)].$$

The global shrinkage parameter κ^g is in $[\kappa, 1]$ and decreases with the correlation ρ , attaining the minimum value, $\kappa^g = \kappa$, when $\rho = 1$ as is clearly seen from (2.12).

The result about the posterior variance follows from Lemma 1. ■

Proof of Proposition 7. The prior joint distribution of the true and estimated alphas is given by the following expression, where we focus on factor 1 without loss of generality:

$$\begin{bmatrix} \alpha^1 \\ \hat{\alpha}^1 \\ \vdots \\ \hat{\alpha}^N \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_c^2 + \tau_w^2 & \tau_c^2 + \tau_w^2 & \tau_c^2 & \cdots & \tau_c^2 \\ \tau_c^2 + \tau_w^2 & \tau_c^2 + \tau_w^2 + \sigma^2/T & & & \tau_c^2 + \rho\sigma^2/T \\ \tau_c^2 & & & & \\ \vdots & & & \ddots & \\ \tau_c^2 & \tau_c^2 + \rho\sigma^2/T & & & \tau_c^2 + \tau_w^2 + \sigma^2/T \end{bmatrix} \right) \quad (2.35)$$

The posterior alpha of factor 1 is therefore normally distributed with a mean derived using the standard formula for conditional normal distributions (2.33):

$$E(\alpha^1 | \hat{\alpha}^1, \dots, \hat{\alpha}^N) = \begin{bmatrix} \tau_c^2 + \tau_w^2 \\ \tau_c^2 \\ \vdots \\ \tau_c^2 \end{bmatrix}^\top \begin{bmatrix} \tau_c^2 + \tau_w^2 + \sigma^2/T & & & \\ & \tau_c^2 + \rho\sigma^2/T & & \\ & & \ddots & \\ \tau_c^2 + \rho\sigma^2/T & & & \tau_c^2 + \tau_w^2 + \sigma^2/T \end{bmatrix}^{-1} \begin{bmatrix} \hat{\alpha}^1 \\ \vdots \\ \hat{\alpha}^N \end{bmatrix}$$

We next use Lemma 2 and its notation, i.e., $a = \tau_c^2 + \tau_w^2 + \sigma^2/T$, $b = \tau_c^2 + \rho\sigma^2/T$, and c', d are defined accordingly, where we use the notation c' to avoid confusion with the c

in equation (2.14). This application of Lemma 2 yields

$$\begin{aligned}
E(\alpha^1 | \hat{\alpha}^1, \dots, \hat{\alpha}^N) &= \begin{bmatrix} \tau_c^2 + \tau_w^2 \\ \tau_c^2 \\ \vdots \\ \tau_c^2 \end{bmatrix}^\top \begin{bmatrix} c' & & d \\ & \ddots & \\ d & & c' \end{bmatrix} \begin{bmatrix} \hat{\alpha}^1 \\ \vdots \\ \hat{\alpha}^N \end{bmatrix} \\
&= \begin{bmatrix} \tau_c^2(c' + d(N-1)) + \tau_w^2 c' \\ \tau_c^2(c' + d(N-1)) + \tau_w^2 d \\ \vdots \\ \tau_c^2(c' + d(N-1)) + \tau_w^2 d \end{bmatrix}^\top \begin{bmatrix} \hat{\alpha}^1 \\ \vdots \\ \hat{\alpha}^N \end{bmatrix} \\
&= (\tau_c^2(c' + d(N-1)) + \tau_w^2 d) N \hat{\alpha}^\cdot + \tau_w^2(c' - d) \hat{\alpha}^1 \\
&= \left(\tau_c^2 \frac{N}{a + b(N-1)} - \tau_w^2 \frac{bN}{(a-b)(a+b(N-1))} \right) \hat{\alpha}^\cdot + \tau_w^2 \frac{1}{a-b} \hat{\alpha}^1 \\
&= \frac{\tau_c^2}{b + \frac{a-b}{N}} \hat{\alpha}^\cdot + \frac{\tau_w^2}{a-b} \left(\hat{\alpha}^1 - \frac{1}{1 + \frac{a-b}{bN}} \hat{\alpha}^\cdot \right) \\
&= \frac{\tau_c^2}{\tau_c^2 + \rho\sigma^2/T + \frac{\tau_w^2 + (1-\rho)\sigma^2/T}{N}} \hat{\alpha}^\cdot + \frac{\tau_w^2}{\tau_w^2 + (1-\rho)\sigma^2/T} \left(\hat{\alpha}^1 - \frac{1}{1 + \frac{\tau_w^2 + (1-\rho)\sigma^2/T}{(\tau_c^2 + \rho\sigma^2/T)N}} \hat{\alpha}^\cdot \right) \\
&= \frac{1}{1 + \frac{\rho\sigma^2}{\tau_c^2 T} + \frac{\tau_w^2 + (1-\rho)\sigma^2/T}{\tau_c^2 N}} \hat{\alpha}^\cdot + \frac{1}{1 + \frac{(1-\rho)\sigma^2}{\tau_w^2 T}} \left(\hat{\alpha}^1 - \frac{1}{1 + \frac{\tau_w^2 + (1-\rho)\sigma^2/T}{(\tau_c^2 + \rho\sigma^2/T)N}} \hat{\alpha}^\cdot \right)
\end{aligned}$$

The posterior has conditional variance

$$\begin{aligned}
\text{Var}(\alpha^1 | \hat{\alpha}^1, \dots, \hat{\alpha}^N) &= \tau_c^2 + \tau_w^2 - \begin{bmatrix} \tau_c^2 + \tau_w^2 \\ \tau_c^2 \\ \vdots \\ \tau_c^2 \end{bmatrix}^\top \begin{bmatrix} c' & & & d \\ & \ddots & & \\ & & \ddots & \\ d & & & c' \end{bmatrix} \begin{bmatrix} \tau_c^2 + \tau_w^2 \\ \tau_c^2 \\ \vdots \\ \tau_c^2 \end{bmatrix} \\
&= \tau_c^2 + \tau_w^2 - \begin{bmatrix} \tau_c^2(c' + d(N-1)) + \tau_w^2 c' \\ \tau_c^2(c' + d(N-1)) + \tau_w^2 d \\ \vdots \\ \tau_c^2(c' + d(N-1)) + \tau_w^2 d \end{bmatrix}^\top \begin{bmatrix} \tau_c^2 + \tau_w^2 \\ \tau_c^2 \\ \vdots \\ \tau_c^2 \end{bmatrix} \\
&= \tau_c^2 + \tau_w^2 - (\tau_c^2(c' + d(N-1)) + \tau_w^2 c')(\tau_c^2 + \tau_w^2) \\
&\quad - (\tau_c^2(c' + d(N-1)) + \tau_w^2 d)\tau_c^2(N-1) \\
&\rightarrow \tau_c^2 + \tau_w^2 - (\tau_c^2(\frac{1}{a-b} - \frac{1}{a-b}) + \tau_w^2 \frac{1}{a-b})(\tau_c^2 + \tau_w^2) \\
&\quad - (\tau_c^2 \frac{1}{b} - \tau_w^2 \frac{1}{a-b})\tau_c^2 \\
&= \tau_c^2 + \tau_w^2 - \left(\tau_w^4 \frac{1}{a-b} + \tau_c^4 \frac{1}{b} \right) \\
&= \tau_c^2 + \tau_w^2 - \left(\frac{\tau_w^4}{\tau_w^2 + (1-\rho)\sigma^2/T} + \frac{\tau_c^4}{\tau_c^2 + \rho\sigma^2/T} \right)
\end{aligned}$$

The last results follow from Lemma 1. ■

Proof of Proposition 8. We write the joint prior distribution of true and observed alphas in the multi-level hierarchical model as

$$\begin{pmatrix} \alpha \\ \hat{\alpha} \end{pmatrix} \sim N \left(\alpha^0 \mathbf{1}_{2NK}, \begin{pmatrix} \Omega & \Omega \\ \Omega & \Omega + \Sigma/T \end{pmatrix} \right) \quad (2.36)$$

The posterior mean vector of true alphas is computed via (2.33):

$$\begin{aligned}
E(\alpha | \hat{\alpha}) &= \mathbf{1}_{NK} \alpha_0 + \Omega (\Omega + \Sigma/T)^{-1} (\hat{\alpha} - \mathbf{1}_{NK} \alpha_0) \\
&= (\Omega^{-1} + T\Sigma^{-1})^{-1} (\Omega^{-1} \mathbf{1}_{NK} \alpha_0 + T\Sigma^{-1} \hat{\alpha})
\end{aligned}$$

using that $(\Omega + \Sigma/T)^{-1} = \Omega^{-1} - \Omega^{-1} (\Omega^{-1} + T\Sigma^{-1})^{-1} \Omega^{-1}$ by the Woodbury matrix identity. The posterior variance is computed similarly via (2.33) and the same application of the Woodbury matrix identity as

$$\text{Var}(\alpha | \hat{\alpha}) = \Omega - \Omega (\Omega + \Sigma/T)^{-1} \Omega = (\Omega^{-1} + T\Sigma^{-1})^{-1}.$$

■

Proof of Proposition 9. Based on the definition of the Bayesian FDR, we have:

$$\begin{aligned}
\text{FDR}^{\text{Bayes}} &= E \left(\frac{\sum_i 1_{\{i \text{ false discovery}\}}}{\sum_i 1_{\{i \text{ discovery}\}}} \middle| \hat{\alpha}^1, \dots, \hat{\alpha}^N, \tau \right) \\
&= \frac{1}{\sum_i 1_{\{i \text{ discovery}\}}} E \left(\sum_i 1_{\{i \text{ false discovery}\}} \middle| \hat{\alpha}^1, \dots, \hat{\alpha}^N, \tau \right) \quad (2.37) \\
&= \frac{1}{\sum_i 1_{\{i \text{ discovery}\}}} \sum_i \text{Pr}(i \text{ false discovery} | \hat{\alpha}^1, \dots, \hat{\alpha}^N, \tau) \\
&= \frac{1}{\#\text{discoveries}} \sum_{i \text{ discovery}} p\text{-null}_i \\
&\leq 2.5\%
\end{aligned}$$

■

Proof of Proposition 10. The posterior mean alpha is

$$\begin{aligned}
E(\alpha | \hat{\alpha}, \hat{\alpha}^{\text{oos}}) &= \begin{bmatrix} \tau^2 & \tau^2 \end{bmatrix} \begin{bmatrix} \tau^2 + \bar{\sigma}_T^2 & \tau^2 \\ \tau^2 & \tau^2 + \sigma_{\text{oos}}^2 \end{bmatrix}^{-1} \begin{bmatrix} \hat{\alpha} - \bar{\varepsilon} \\ \hat{\alpha}^{\text{oos}} \end{bmatrix} \\
&= \frac{1}{\det} \begin{bmatrix} \tau^2 & \tau^2 \end{bmatrix} \begin{bmatrix} \tau^2 + \sigma_{\text{oos}}^2 & -\tau^2 \\ -\tau^2 & \tau^2 + \bar{\sigma}_T^2 \end{bmatrix} \begin{bmatrix} \hat{\alpha} - \bar{\varepsilon} \\ \hat{\alpha}^{\text{oos}} \end{bmatrix} \\
&= \frac{\tau^2}{\det} (\sigma_{\text{oos}}^2 (\hat{\alpha} - \bar{\varepsilon}) + \bar{\sigma}_T^2 \hat{\alpha}^{\text{oos}}) \\
&= \frac{\tau^2 (\bar{\sigma}_T^2 + \sigma_{\text{oos}}^2)}{\tau^2 (\bar{\sigma}_T^2 + \sigma_{\text{oos}}^2) + \bar{\sigma}_T^2 \sigma_{\text{oos}}^2} (w(\hat{\alpha} - \bar{\varepsilon}) + (1 - w)\alpha^{\text{oos}}) \\
&= \frac{\tau^2}{\tau^2 + \bar{\sigma}_T^2 \sigma_{\text{oos}}^2 / (\bar{\sigma}_T^2 + \sigma_{\text{oos}}^2)} (w(\hat{\alpha} - \bar{\varepsilon}) + (1 - w)\alpha^{\text{oos}}) \\
&= \frac{1}{1 + \frac{1}{\tau^2 (\bar{\sigma}_T^2 + \sigma_{\text{oos}}^2)}} (w(\hat{\alpha} - \bar{\varepsilon}) + (1 - w)\alpha^{\text{oos}})
\end{aligned}$$

using the notation $\bar{\sigma}_T^2 = \bar{\sigma}^2/T$, $\sigma_{\text{oos}}^2 = \sigma^2/T^{\text{oos}}$, and

$$\det = (\tau^2 + \bar{\sigma}_T^2)(\tau^2 + \sigma_{\text{oos}}^2) - \tau^4 = \tau^2(\bar{\sigma}_T^2 + \sigma_{\text{oos}}^2) + \bar{\sigma}_T^2 \sigma_{\text{oos}}^2.$$

■

5.3 Empirical Bayes Estimation

For convenient reference, we restate the multi-level hierarchical model of Section 1. For a factor i in cluster j and corresponding to signal n , the factor is

$$f_t^i = \alpha^i + \beta^i r_t^m + \varepsilon_t^i$$

with

$$\alpha^i = \alpha^o + c^j + s^n + w^i$$

where the alpha components are $\alpha^o = 0$, $c^j \sim N(0, \tau_c^2)$, $s^n \sim N(0, \tau_s^2)$, and $w^i \sim N(0, \tau_w^2)$. We write alpha in vector form as

$$\alpha = \alpha^o 1_{NK} + Mc + Zs + w \tag{2.38}$$

where $\alpha = (\alpha^1, \dots, \alpha^{NK})'$, $c = (c^1, \dots, c^J)'$, $s = (s^1, \dots, s^N)'$, $w = (w^1, \dots, w^{NK})'$, M is the $NK \times J$ matrix of cluster memberships, and Z is the $NK \times N$ matrix indicating the characteristic that factor i is based on. Given the hyperparameters $(\alpha^o, \tau_c, \tau_s, \tau_w)$, the prior mean and covariance matrix of alphas are

$$E[\alpha] = 0, \quad \Omega \equiv \text{Var}(\alpha) = MM'\tau_c^2 + ZZ'\tau_s^2 + I_{NK}\tau_w^2. \tag{2.39}$$

The vector of return shocks is $\varepsilon_t = (\varepsilon_t^1, \dots, \varepsilon_t^{NK})'$ which is distributed $\varepsilon_t \sim N(0, \Sigma)$.

Given this structure, we estimate the model as follows. The vector of factor returns $f_t = (f_t^1, \dots, f_t^{NK})'$ has marginal likelihood—that is, after integrating out the uncertain alpha components—that is distributed as

$$f_t \sim N(0, [\Omega + \Sigma])$$

or, equivalently (treating CAPM betas as known), the estimated alphas are distributed⁴⁷

$$\hat{\alpha} \sim N(0, [\Omega + \Sigma/T]).$$

The matrices Z and M are given by the factor definition and cluster assignment (Table AIV), respectively. We use a plug-in estimate of the factor CAPM-residual return covariance matrix, denoted $\hat{\Sigma}$ (discussed below). Finally, given $\hat{\Sigma}$, Z , and M , we estimate the hyperparameters of the prior distribution, (τ_c, τ_s, τ_w) via MLE based on the marginal likelihood.

This estimation approach is an example of the empirical Bayes method. It approx-

⁴⁷We abstract from uncertainty in CAPM betas to emphasize the Bayesian updating of alphas. Our conclusions are qualitatively insensitive to accounting for beta uncertainty.

imates the fully Bayesian posterior calculation (which requires integrating over a hyperprior distribution of hyperparameters, usually an onerous calculation) by setting the hyperparameters to their most likely values based on the marginal likelihood. It is particularly well suited to hierarchical Bayesian models in which parameters for individual observations share some common structure, so that the realized heterogeneity across individual is informative about sensible values for the hyperparameters of the prior. Our model and estimation approach implementation is a minor variation on Bayesian hierarchical normal mean models that are common in Bayesian statistics (textbook treatments include Efron, 2012; Gelman et al., 2013; Maritz, 2018). We conduct sensitivity analysis to ensure that our results are robust to a wide range of hyperparameters (see Figure A6). Also, we note that our EB methodology is more easily replicable than a full-Bayesian setting with additional hyperpriors as EB relies on a closed-form Bayesian updating rather than a numerical integration.

To ensure cross-sectional stationarity, we scale each factor such that their monthly idiosyncratic volatility is $10\%/\sqrt{12}$ (i.e., 10% annualized). To construct a plug-in estimate of the factor residual return covariance matrix, denoted $\hat{\Sigma}$, we face two main empirical challenges. First, the sample covariance is poorly behaved due the relatively large number of factors compared to the number of time series observations. Second, we have an unbalanced panel because different factors come online at different points in time. To address the first challenge, we impose a block equicorrelation structure on Σ based on factors' cluster membership.⁴⁸ The correlation between factors in clusters i and j is estimated as the average correlation among all pairs such that one factor is in cluster i and the other is in j . In our global analyses, blocks correspond to region-cluster pairs. To address unbalancedness, we use the bootstrap. In particular, we generate 10,000 bootstrap samples that resample rows of the unbalanced factor return dataset. Each bootstrap sample is, therefore, also unbalanced, and we use this to produce a distribution of alpha estimates. From this we calculate $\hat{\Sigma}/T$ as the covariance of alphas across bootstrap samples (imposing the block equicorrelation structure).

Table AI shows the estimated hyperparameters across different samples. While most of our analysis is based on these full-sample estimates, we also consider rolling-estimates of when considering out-of-sample evidence as seen in Figure A5.

⁴⁸As advocated by Engle and Kelly (2012) and Elton et al. (2006), block equicorrelation constrains all pairs of factors in the same block to share a single correlation parameter, and likewise for cross-block correlations. This stabilizes covariance matrix estimates by dramatically reducing the parameterization of the correlation matrix, while leaving the individual variance estimates unchanged.

Sample	τ_c	τ_w	τ_s
USA	0.35%	0.21%	
Developed	0.24%	0.18%	
Emerging	0.32%	0.24%	
USA, Developed & Emerging	0.30%	0.19%	0.10%
World	0.37%	0.23%	
World ex. US	0.29%	0.20%	
USA - Mega	0.26%	0.16%	
USA - Large	0.31%	0.18%	
USA - Small	0.44%	0.26%	
USA - Micro	0.48%	0.32%	
USA - Nano	0.42%	0.28%	

Table AI. Hyperparameters of the prior distribution estimated by maximum likelihood. Here, τ_c is the estimated dispersion in cluster alphas (e.g., the dispersion in the alpha of the value cluster alpha, momentum cluster, and so on). When we estimate a single region, τ_w is the idiosyncratic dispersion of alphas within each cluster. When we jointly estimate several regions, then τ_s is the estimated dispersion in alphas across signals within each cluster, and τ_w is the estimated idiosyncratic dispersion in alphas for factors identified by their signal and region.

5.4 Differences in Sample and Factor Construction

Here we provide further details on the difference in sample and factor construction vs. Hou et al. (2020b) accounting for the difference between the first two bars in Figure 1 in our introduction. To re-iterate, with raw returns and capped value weights, we find a replication rate of 55.6% where as Hou et al. (2020b) finds a replication rate of 35%.

This difference has the following decomposition.⁴⁹ First, Hou et al. (2020b) focus their analysis on value-weighted factors rather than the standard Fama and French (1993b) methodology that gives half the weight to small stocks (or equal-weighting that gives even more weight to small stocks). However, pure value weighting sometimes leads to excessively concentrated portfolios that mask the behavior of factors.⁵⁰ We use a weighting scheme that we refer to as “capped value-weighting” that winsorizes market caps at the NYSE 80th percentile. This weighting is a helpful compromise between pure value-weighting and the Fama-French method since our factors continue to emphasize large stocks, but the capped scheme avoids undue skewness toward a few mega stocks, which in turn produces more robust factor behavior over time and across countries. Capped value weights contribute +9.2% to our higher replication rate. Figure A1 reproduces Figure 1 with straight value weights.

⁴⁹Note that the attribution to specific changes depends on the order in which the changes are applied.

⁵⁰For example, Nokia stock accounted for more than 70% of the total market capitalization in Finland in 1999 and 2000.

Second, for each characteristic, Hou et al. (2020b) construct three variations on each factor having either 1-month, 6-month, or 12-month holding periods. They treat these as separate factors so that their factor count essentially multiplies their characteristics count by a factor of three. In contrast, we focus on 1-month returns because this is the horizon of interest in almost all of the original studies (and we believe it is the most economically meaningful since it uses the most current data as theory dictates). Our focus on only the 1-month holding period factor for each characteristic contributes +5.0% to our replication rate.

Third, we use a longer sample, which contributes +8.3% to the difference in replication rate. Fourth, we add 15 factors to our sample that were previously studied in the literature but not studied by Hou et al. (2020b), which has a no effect on the replication rate.

Finally, we use tercile spreads and breakpoints based on all stocks above the NYSE 20th percentile (i.e., non-micro-caps), while Hou et al. (2020b) use decile spreads and breakpoints based on all NYSE stocks. Our more conservative method leads to a -6.0% drop in the replication rate. The remaining +4.1% difference in replication rates is due to minor construction and sample details⁵¹. We discuss this decomposition further in Section 2, where we detail our factor construction choices and discuss why we prefer them.

Replication Rate with Uncapped Value Weights

In Figure A1, we show an alternative version of Figure 1 with factors constructed using straight (as opposed to capped) value weights. It shows that all of main our conclusions remain similar. Our ultimate replication rate in this case is 79.8% (based on global data and Bayesian model estimates).

⁵¹For example we always lag accounting data four months, they use a mixture of updating schemes and our set of factors is not identical to that in Hou et al. (2020b).

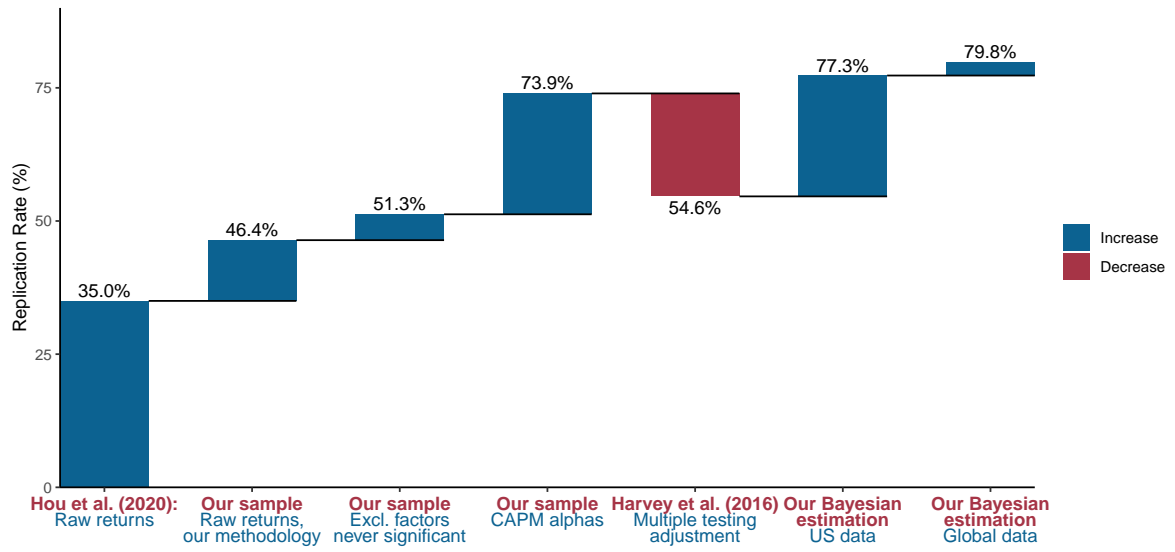


Figure A1. Replication Rates Versus the Literature (Uncapped Value-weighting)

Note: This figure reproduces the analysis of figure 1 using uncapped value weights to construct factors.

5.5 Additional Time-Series Results

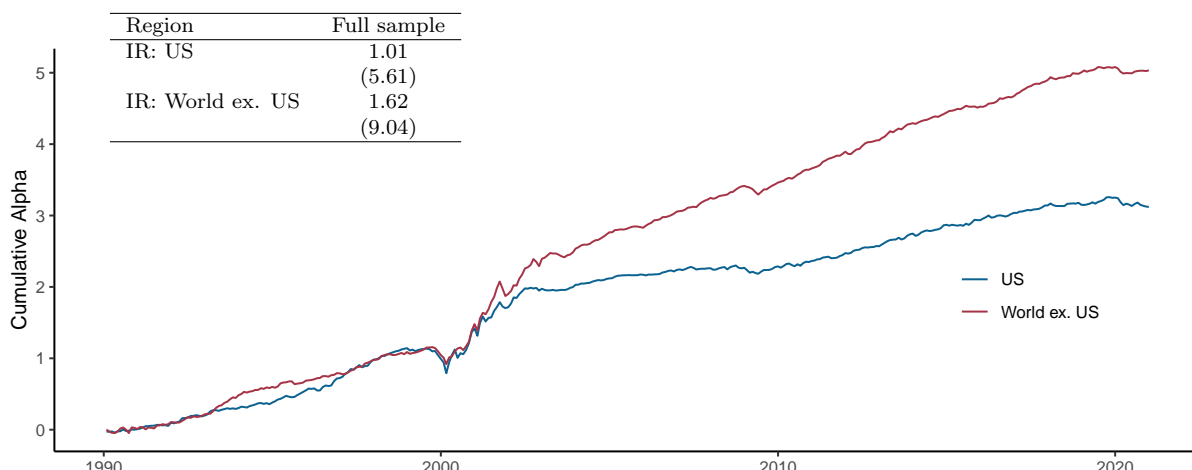


Figure A2. Out-of-sample performance of significant factors under empirical Bayes

Note: The figure shows the cumulative CAPM alpha of an average of factors significant under our empirical Bayes framework. The significance cutoffs are re-estimated each year with the available data. Factors are eligible for inclusion after the sample period in the original paper, so all returns are out-of-sample. The table shows the information ratio (alpha divided by residual volatility) for the full sample (1990-2020) with t -statistics in parentheses.

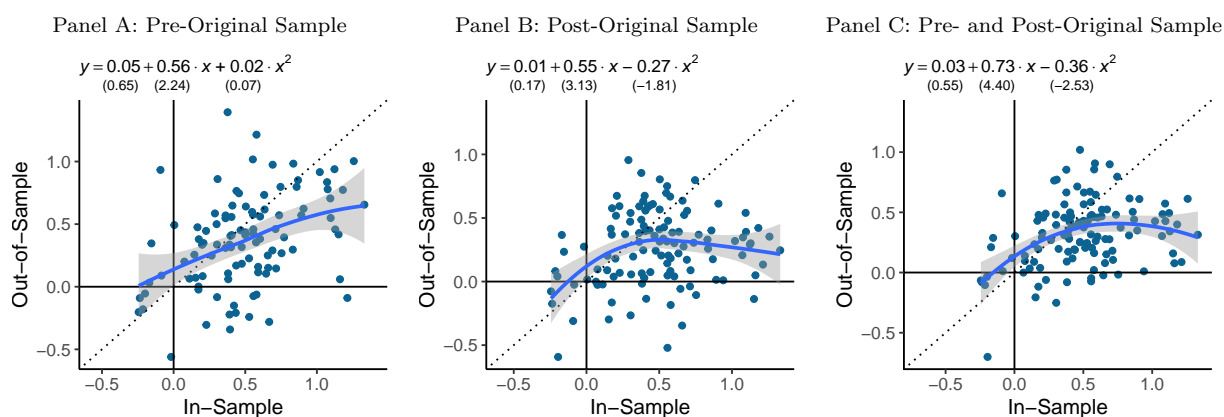


Figure A3. In-Sample versus Out-of-Sample Alphas for US Factors

Note: The figure plots OLS alphas for US factors during the in-sample period (i.e., the period studied in the original publication) versus out-of-sample alphas. In Panel A, out-of-sample is the time period before the in-sample period. In Panel B, out-of-sample is the time period before the in-sample period. In Panel C, out-of-sample includes both the time period before and after the in-sample period. We require at least five years of out-of-sample data for a factor to be included, amounting to 102, 115 and 119 factors in panel A, B and C. The figure also reports feasible GLS estimates of out-of-sample alphas on in-sample alphas and in-sample alphas squared. To implement feasible GLS, we assume that the error variance-covariance matrix is proportional to the full-sample CAPM residual variance-covariance matrix, $\hat{\Sigma}/T$, described in section 5.3. The blue line is a local polynomial regression fit where observations are weighted by their vicinity to the point on the x-axis. The shaded area is 95% confidence bands. The dotted line is the 45° line.

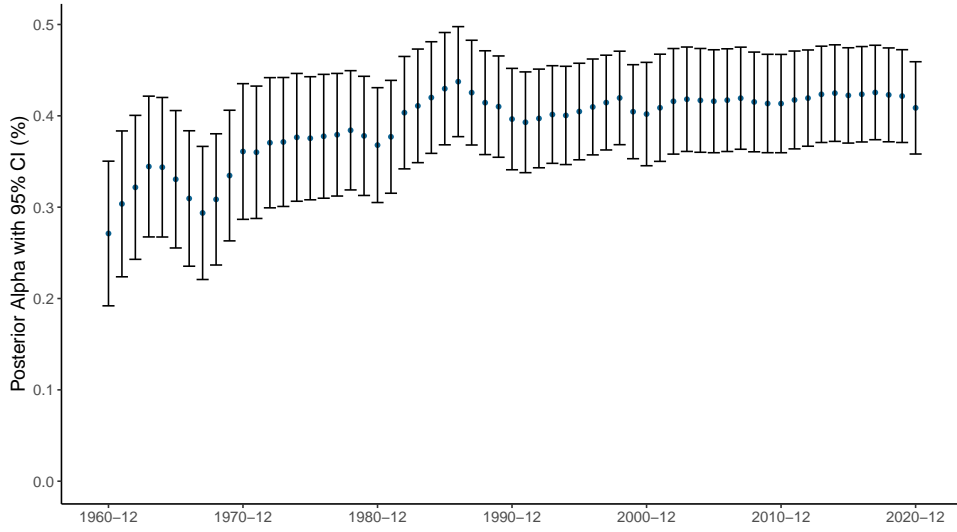


Figure A4. World Factor Alpha Posterior Distribution Over Time

Note: The figure reports the CAPM alpha and 95% posterior confidence interval for an equal-weighted portfolio of World factors based on EB posteriors re-estimated in December each year. In contrast to figure 9, we re-estimate τ_c and τ_w at each point in time. Figure A5 shows how the estimated taus evolve over time.

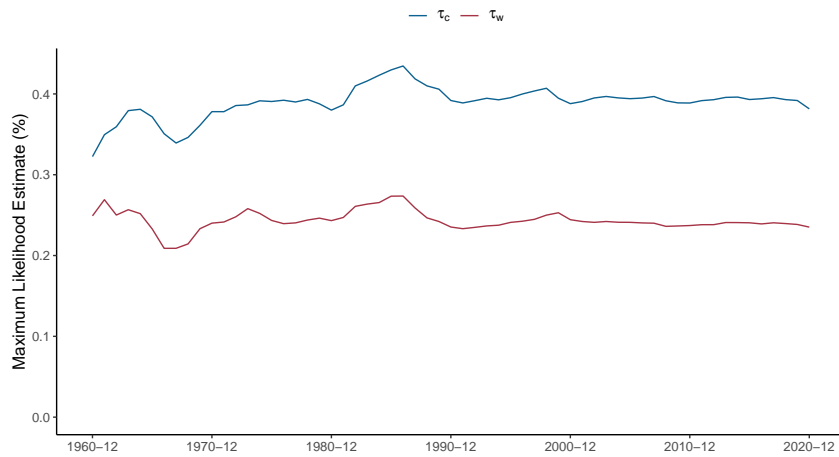


Figure A5. World Factor Hyperparameters Over Time

Note: The figure reports the τ_c and τ_w used in figure A4.

5.6 Economic Benefit of More Powerful Tests

Table AII. The Economic Benefit of More Powerful Tests

	Region		
	US	Developed ex. US	Emerging
	(1)	(2)	(3)
Alpha	0.35*** (5.05)	0.24*** (5.33)	0.27*** (3.66)
Market Beta	-0.12*** (-4.33)	-0.09*** (-5.68)	-0.04*** (-3.14)
Observations	540	420	388
Adjusted R^2	0.17	0.18	0.03

Note: The dependent variable is an equal-weighted portfolio of factors that are significant under empirical Bayes (EB), but not under OLS with the Benjamini-Yekutieli adjustment (BY). A factor is significant under EB when the probability of a negative alpha is below 2.5%. A factor is significant under BY when the adjusted two-sided p -value is below 5%, and the OLS alpha estimate is positive. Starting in 1959, we update the posterior distribution and the OLS estimates by the end of each year and invest in marginally significant factors over the subsequent year. To avoid lookahead bias, we only use factors after the sample in the original paper has ended. We only consider factors found to be significant by the original reference. The alpha estimates are in percentages, with t -statistics in parentheses. Standard errors are computed following Newey and West (1987) with 6 lags. The stars indicate * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

5.7 Accounting for Publication Bias

Harvey et al. (2016b) provides a framework to estimate the total number of factors researchers have tried. The framework is based on t -statistics of published factors and estimation framework to determine the number of unobserved factors.

One set of simulations is constructed to match the baseline scenario of (Harvey et al., 2016b, Table 5.A, row 1), which estimates that researchers have tried $M = 1,297$ factors, of which 39.6% of have zero alpha and another is based on the more conservative scenario of (Harvey et al., 2016b, Table 5.B, row 1), which implies that researchers have tried 2458 factors, of which 68.3% have zero alpha. Harvey et al. (2016b) states that “the average annual Sharpe ratio for these [true] factors is 0.44.”

To incorporate these unobserved factors into our framework, we proceed as follows for the baseline scenario. We simulate a total of 1,300 factors in 26 clusters of 50 factors per cluster. We let all factors in 10 clusters have true alphas equal to zero while the remaining clusters have non-zero true alphas. For each of the clusters with non-zero alphas, we set the cluster alpha to $c^j = 0.44 \times 10\% / 12$ so that the monthly abnormal return corresponds

to an annual Sharpe ratio of 0.44 given the annual volatility of 10%. Finally, we draw each factor's true alpha from $\alpha^i \sim N(c^j, \tau_w^2)$, and then simulate 70 years of monthly returns with within-cluster correlation of 0.58 and 0.02 otherwise.⁵² Finally, we estimate prior parameters τ using this data with the same method that we used on the observed data. We repeat this simulation process and compute the average τ_c , which is interpreted as a value that accounts for unobserved factors of the form implied by Harvey et al. (2016b). We note that we are implicitly assuming that the unobserved factors belong to different clusters, such that observing new poor performing factors would lead to more shrinkage toward zero via a lower τ_c , but not via different cluster mean returns.

Similarly for the conservative scenario, we simulate a total of 2500 factors in 50 clusters of 50 factors per cluster. We let all factors in 16 clusters have true alphas equal to zero while the remaining clusters have non-zero true alphas as described above. Figure A6 shows the US replication rate under these alternative hyperparameters of the prior distribution.

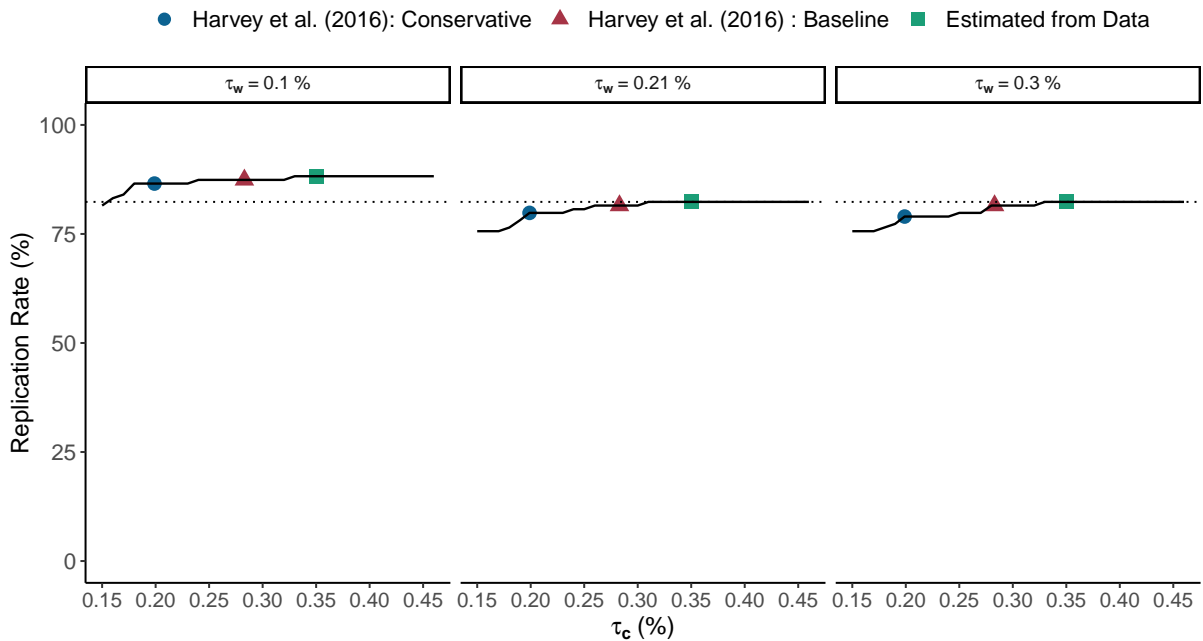


Figure A6. Replication Rate with Prior Estimated in Light of Unobserved Factors

Note: The figure shows how the replication rate in the US varies when changing the τ_c and τ_w parameter. The dotted line shows our replication rate of 82.4%. The data estimate of τ_w is 0.21%. The green square, highlights the value estimated in the data $\tau_c = 0.35\%$. The red triangle and the blue circle highlights values that are found by estimating the empirical Bayes model according to assumptions about unobserved factors from Harvey et al. (2016b). The values are $\tau_c = 0.28\%$ in the baseline scenario and $\tau_c = 0.20\%$ in the conservative scenario. A description of this approach can be found in the appendix, section 5.7.

⁵²The values are calibrated to match the data on US factors.

5.8 Results by Cluster, Region, and Size

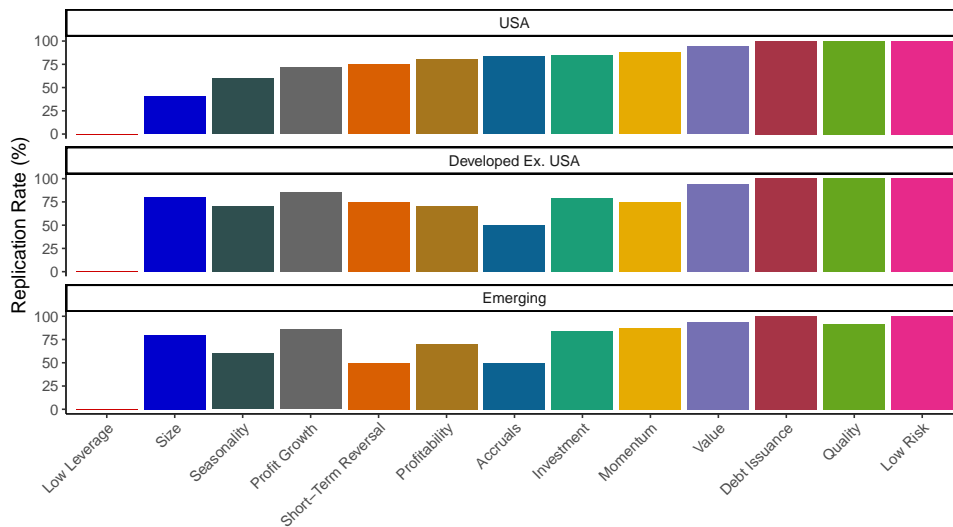


Figure A7. Replication Rates across Regions by Cluster

Note: Share of factors within each cluster where the 95% posterior intervals does not include zero.

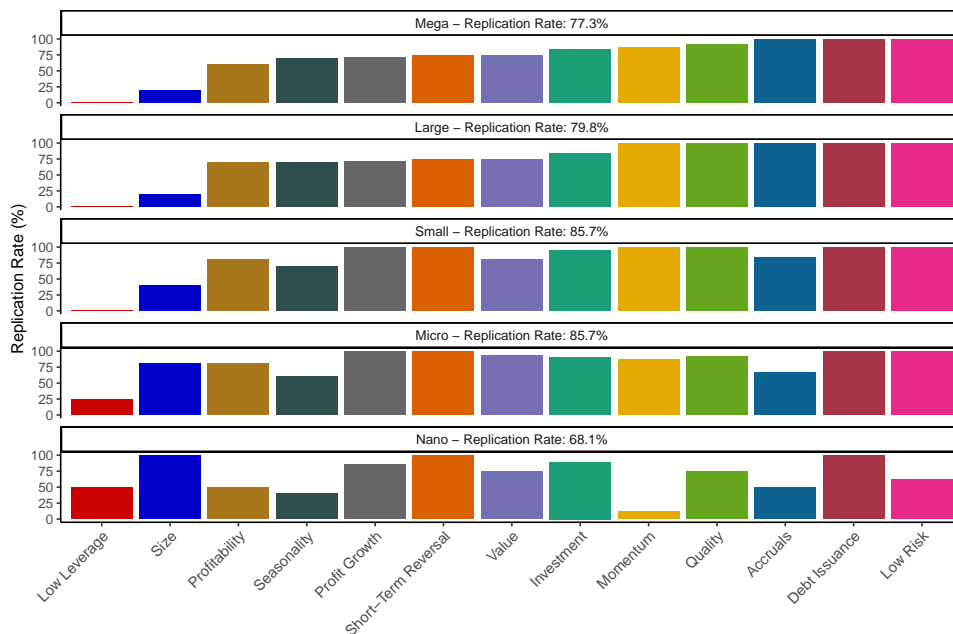


Figure A8. Replication Rates across Size Groups by Cluster

Note: The figure shows replication rates for US factors created within a size group using rank weights. Mega stocks have a market cap higher than the 80th percentile of NYSE stocks, large stocks are between the 80th and 50th percentile, small stocks are between the 50th and 20th percentile, micro stocks are between the 20th and 1st percentile and nano stocks have a market cap below the 1st percentile of NYSE stocks.

5.9 Further Results on the Tangency Portfolio

In this section, we elaborate on the influence of factors on the tangency portfolio (TPF). Figure A9 and A10 shows the tangency weights across regions and size groups. Most notably, the size cluster is much more important outside of the US and among smaller stocks.

The previous TPF analysis, and the results shown in figure 13 has used the 13 cluster level portfolios in addition to the market portfolio as inputs. In the remainder of this section, we build the TPF at the factor level by using the 119 US factors that were found to be significant by an earlier paper. We start the analysis in 1972 to ensure that all factors have non-missing data. The main issue with a factor level analysis, is estimating the covariance matrix. We follow Pedersen (2021) and adjust the covariance matrix by shrinking the correlations towards zero

$$\Sigma_w = \sigma[(1 - w)\mathbf{\Omega} + w\mathbf{I}]\sigma$$

where $\mathbf{\Omega}$ is the sample correlation matrix, σ is a matrix with the sample volatilities on the diagonal and zero elsewhere, \mathbf{I} is the identity matrix and w is a shrinkage parameter. The tangency weights are recovered from the standard formula on the adjusted covariance matrix. This approach requires choosing the shrinkage parameter. Ideally, we want to choose the amount of shrinkage to maximize out-of-sample Sharpe ratio. We implement this intuition via. five-fold cross validation. In each fold, we estimate the tangency weights with a given shrinkage parameter using 4/5 of the data, and compute the realized Sharpe ratio on the remaining 1/5. We repeat this procedure 5 times, and compute the average realized Sharpe ratio for $w \in (0, 0.1, \dots, 1)$. In unreported results, we find that the optimal shrinkage parameter is $w = 0.5$ ⁵³.

Figure A11 shows the in-sample Sharpe ratio of the tangency portfolio that are allowed to invest in the market portfolio and factors from one clusters. The dashed line shows the Sharpe ratio of the market portfolio. Figure A12 shows the in-sample Sharpe ratio attainable after excluding factors from one cluster at a time. Figure A13 shows the importance of each factor for the cluster TPF. Specifically, we report the drop in the maximal attainable Sharpe ratio within a cluster after excluding one of the cluster factors. Finally, figure A14 shows the importance of each factor for the TPF that includes all factors. Specifically, we eliminate each factor one at a time and record the resulting drop in the in-sample Sharpe ratio.

⁵³The average monthly out-of-sample Sharpe ratio with $w = 0.5$ is 0.56 compared to 0.43 from the unconstrained solution ($w = 0$).

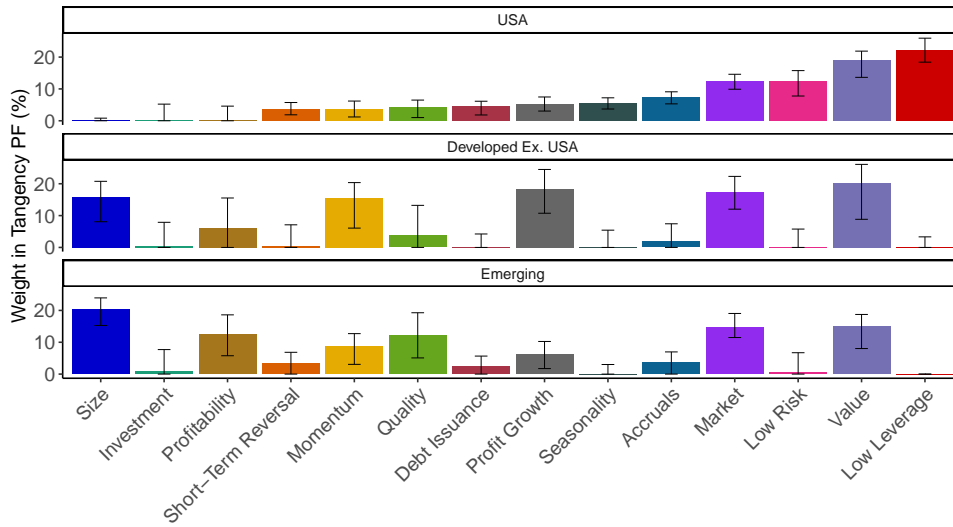


Figure A9. Tangency Portfolio Weights across Regions

Note: Within each region, we compute the cluster return as the equal weighted return of all factors with data available at a given point in time. We further add the regional market return. We estimate the tangency weights following the method of Britten-Jones (1999) with a non-negativity constraint. The error bars are the 90% confidence intervals based on 10,000 bootstrap samples and the percentile method. The data starts in 1952 for the US, 1987 for Developed ex. US and 1994 for Emerging.

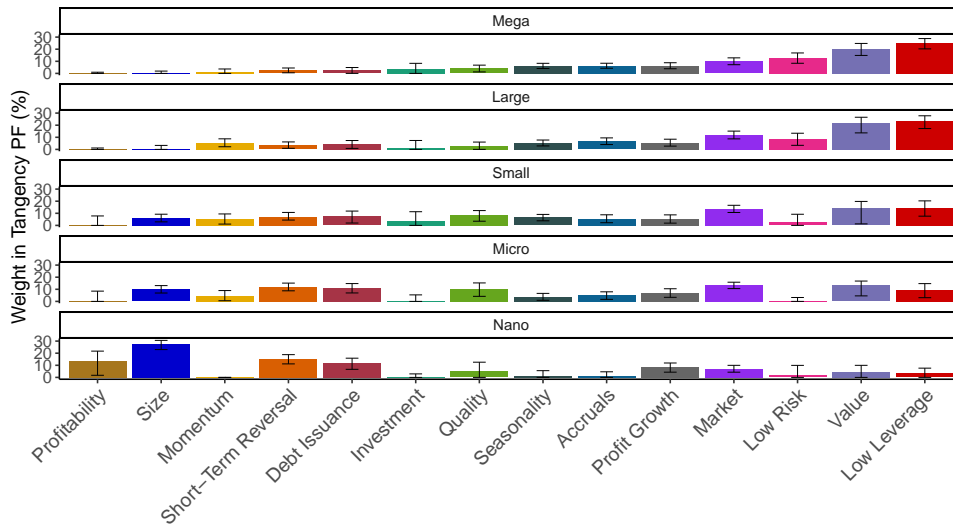


Figure A10. Tangency Portfolio Weights across Size Groups

Note: Within each size group, we compute the cluster return as the equal weighted return of all factors with data available at a given point in time. We only use US data. We add the US market return. We estimate the tangency weights following the method of Britten-Jones (1999) with a non-negativity constraint. The error bars are the 90% confidence intervals based on 10,000 bootstrap samples and the percentile method. The data starts in 1963.

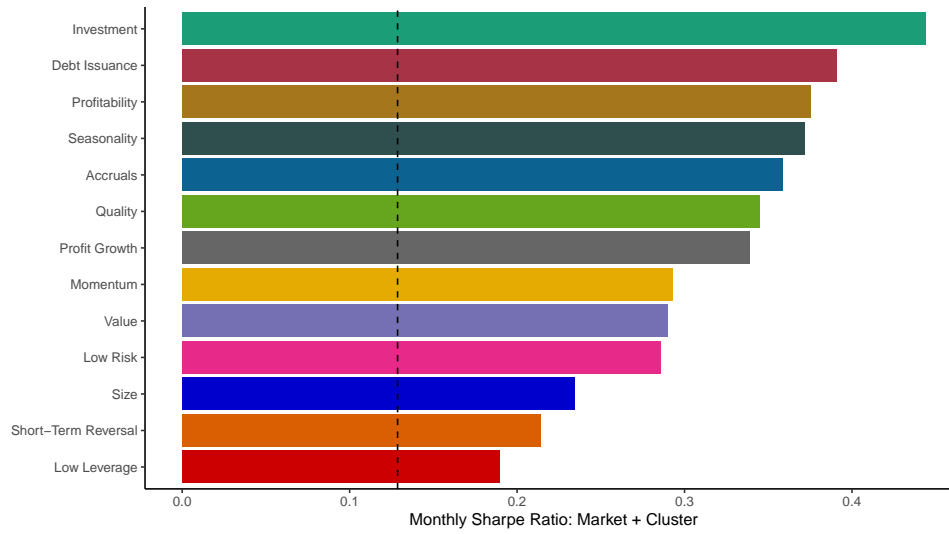


Figure A11. Market + Cluster

Note: Each bar shows the monthly in-sample Sharpe ratio of a tangency portfolio that is allowed to invest in all factor from one cluster plus the market portfolio. We use the simple enhanced portfolio optimization method from Pedersen (2021), with a shrinkage parameter of $w = 0.5$. The analysis is done on US factors from 1972 to 2020.

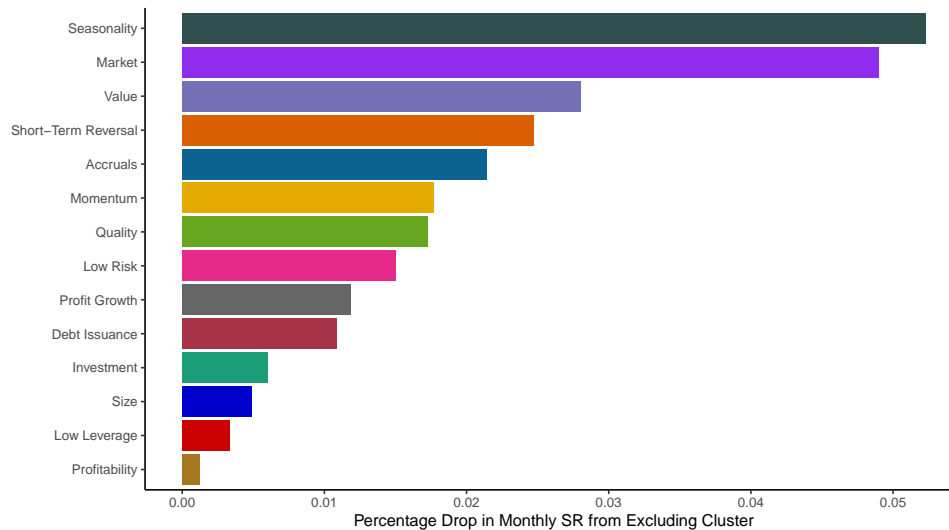


Figure A12. Excluding One Cluster

Note: Each bar shows the monthly in-sample Sharpe ratio of a tangency portfolio that is allowed to invest in the market portfolio and factors from all clusters except one. We use the simple enhanced portfolio optimization method from Pedersen (2021), with a shrinkage parameter of $w = 0.5$. The analysis is done on US factors from 1972 to 2020.

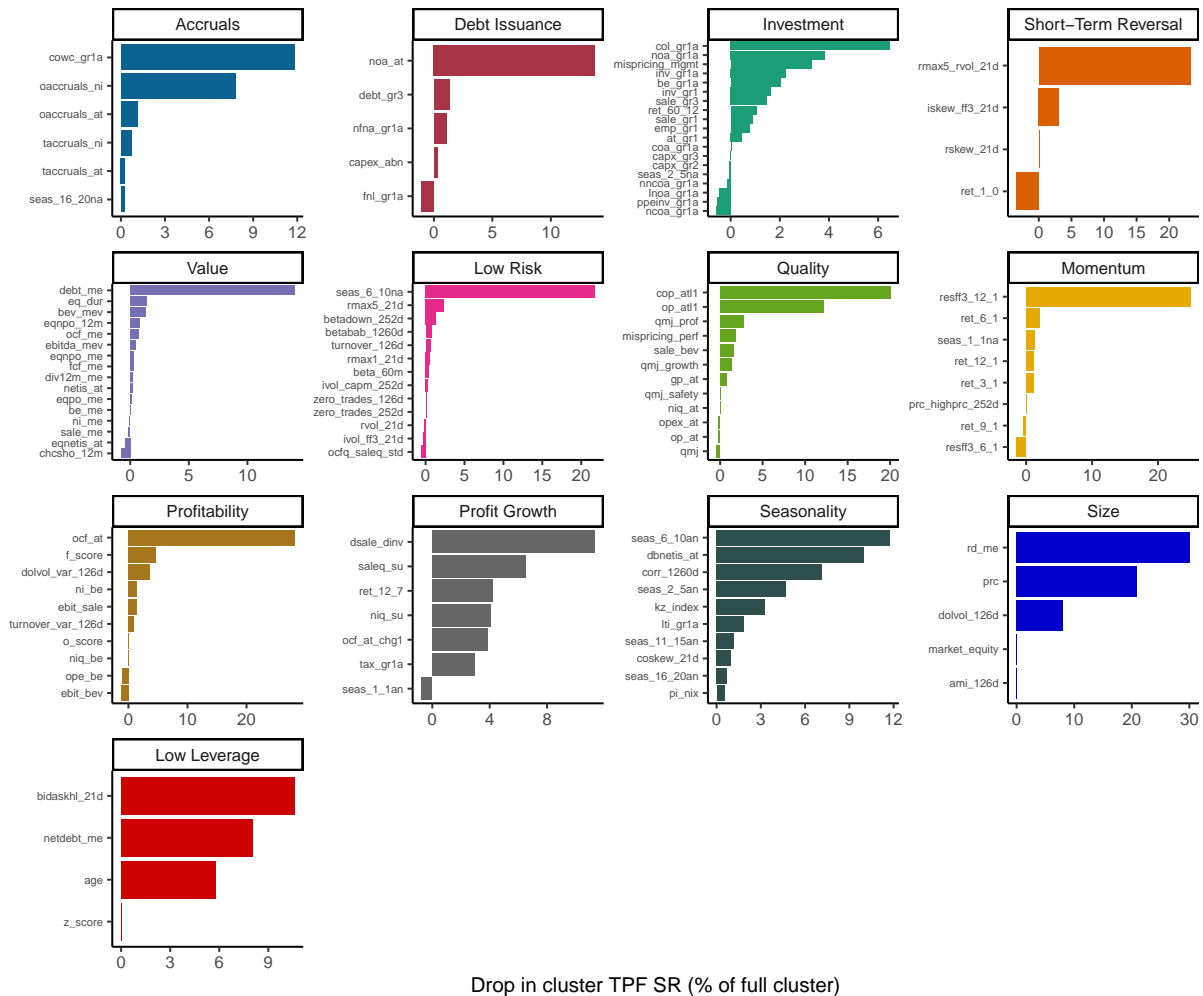


Figure A13. Factor importance for cluster TPF

Note: Each bar shows the difference in the monthly in-sample Sharpe ratio of a tangency portfolio that invest in all factors within a cluster and a tangency portfolio that invest in all cluster factors except one. We show individual factors by their cluster. We use the simple enhanced portfolio optimization method from Pedersen (2021), with a shrinkage parameter of $w = 0.5$. The analysis is done on US factors from 1972 to 2020.

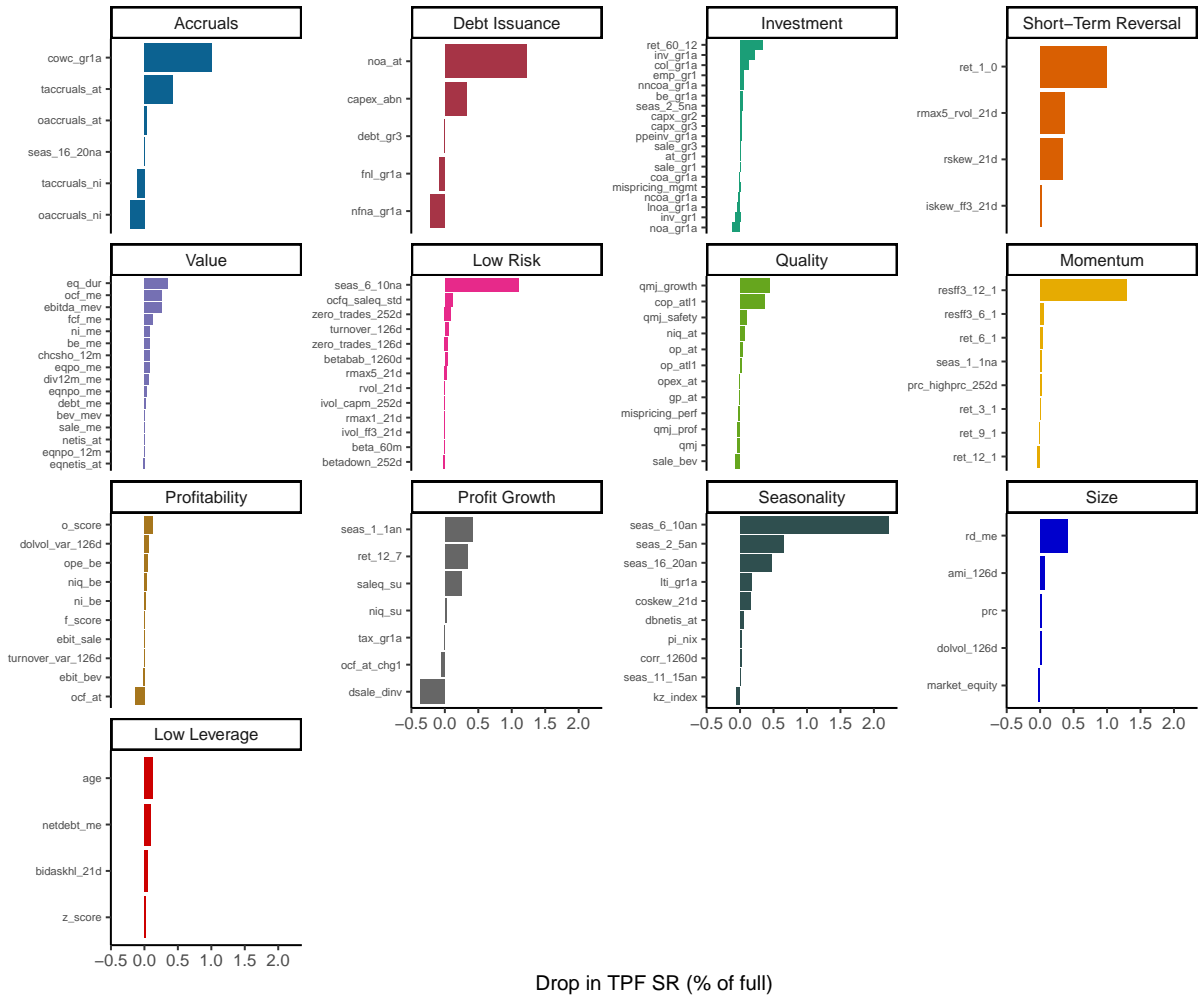


Figure A14. Factor importance for full TPF

Note: Each bar shows the difference in the monthly in-sample Sharpe ratio of a tangency portfolio that invest in all factors and a tangency portfolio that invest in all factor except one. We show individual factors by their cluster. We use the simple enhanced portfolio optimization method from Pedersen (2021), with a shrinkage parameter of $w = 0.5$. The analysis is done on US factors from 1972 to 2020.

5.10 Cluster Construction

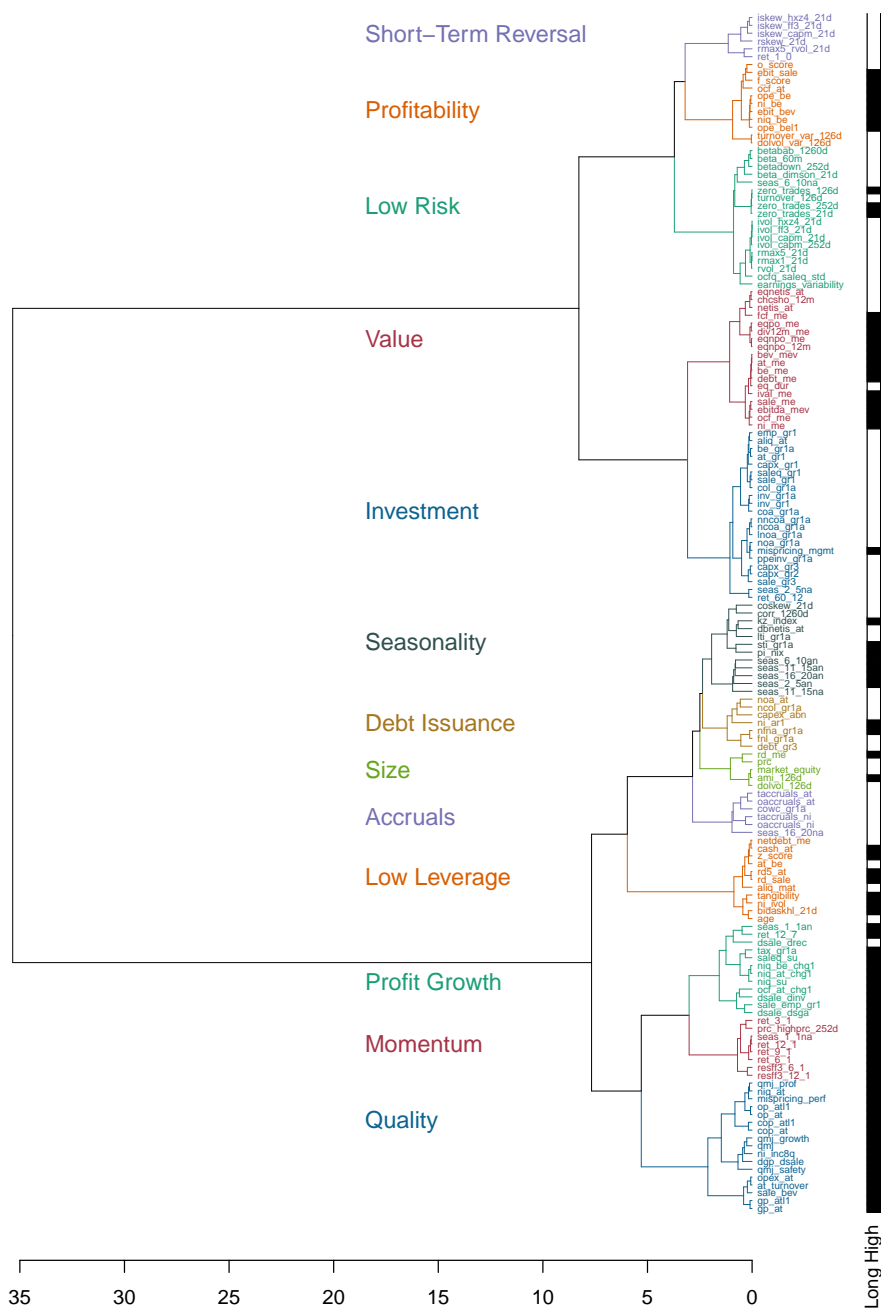


Figure A15. Clustering Factors into Themes

Note: This figure shows a hierarchical clustering of all factors into 13 themes using the sample of US stocks from 1975-2020. Long high indicates whether the factor is long stocks with a high value of the underlying characteristic.

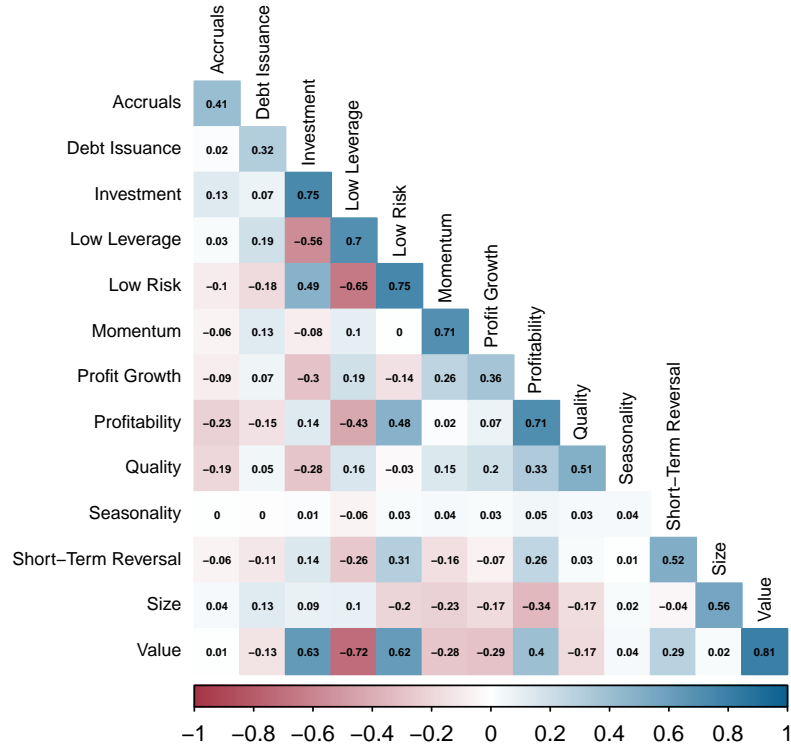


Figure A16. Factor Theme Correlations

Note: This figure shows the average pairwise Pearson correlation between factors from different clusters (off diagonal elements) or between factors in the same cluster (diagonal elements), using data on US stocks during the 1975-2020 period.

5.11 Details on Clusters, Factors, and Countries

Table AIII. Factor and Cluster Details

Description	Variable Name	Citation	Orig. Sample	Orig. Sign	Orig. Signif.
Accruals					
Change in current operating working capital	cowc_gr1a	Richardson et al. (2005)	1962-2001	-1	1
Operating accruals	oaccruals_at	Sloan (1996)	1962-1991	-1	1
Percent operating accruals	oaccruals_ni	Hafzalla Lundholm and Van Winkle (2011)	1989-2008	-1	1
Years 16-20 lagged returns, nonannual	seas_16_20na	Heston and Sadka (2008)	1965-2002	-1	1
Total accruals	taccruals_at	Richardson et al. (2005)	1962-2001	-1	1

Percent total accruals	taccruals_ni	Hafzalla Lundholm and Van Winkle (2011)	1989-2008	-1	1
------------------------	--------------	---	-----------	----	---

Debt Issuance

Abnormal corporate investment	capex_abn	Titman Wei and Xie (2004)	1973-1996	-1	1
Growth in book debt (3 years)	debt_gr3	Lyandres Sun and Zhang (2008)	1970-2005	-1	1
Change in financial liabilities	fnl_gr1a	Richardson et al. (2005)	1962-2001	-1	1
Change in noncurrent operating liabilities	ncol_gr1a	Richardson et al. (2005)	1962-2001	-1	0
Change in net financial assets	nfna_gr1a	Richardson et al. (2005)	1962-2001	1	1
Earnings persistence	ni_ar1	Francis et al. (2004)	1975-2001	1	0
Net operating assets	noa_at	Hirshleifer et al. (2004)	1964-2002	-1	1

Investment

Liquidity of book assets	aliq_at	Ortiz-Molina and Phillips (2014)	1984-2006	-1	0
Asset Growth	at_gr1	Cooper Gulen and Schill (2008)	1968-2003	-1	1
Change in common equity	be_gr1a	Richardson et al. (2005)	1962-2001	-1	1
CAPEX growth (1 year)	capx_gr1	Xie (2001)	1971-1992	-1	0
CAPEX growth (2 years)	capx_gr2	Anderson and Garcia-Feijoo (2006)	1976-1998	-1	1
CAPEX growth (3 years)	capx_gr3	Anderson and Garcia-Feijoo (2006)	1976-1998	-1	1
Change in current operating assets	coa_gr1a	Richardson et al. (2005)	1962-2001	-1	1
Change in current operating liabilities	col_gr1a	Richardson et al. (2005)	1962-2001	-1	1
Hiring rate	emp_gr1	Belo Lin and Bazdresch (2014)	1965-2010	-1	1
Inventory growth	inv_gr1	Belo and Lin (2011)	1965-2009	-1	1
Inventory change	inv_gr1a	Thomas and Zhang (2002)	1970-1997	-1	1
Change in long-term net operating assets	lnoa_gr1a	Fairfield Whisenant and Yohn (2003)	1964-1993	-1	1
Mispricing factor: Management	mispricing_mgmt	Stambaugh and Yuan (2016)	1967-2013	1	1
Change in noncurrent operating assets	ncoa_gr1a	Richardson et al. (2005)	1962-2001	-1	1
Change in net noncurrent operating assets	nncoa_gr1a	Richardson et al. (2005)	1962-2001	-1	1
Change in net operating assets	noa_gr1a	Hirshleifer et al. (2004)	1964-2002	-1	1
Change PPE and Inventory	ppeinv_gr1a	Lyandres Sun and Zhang (2008)	1970-2005	-1	1
Long-term reversal	ret_60-12	De Bondt and Thaler (1985)	1926-1982	-1	1
Sales Growth (1 year)	sale_gr1	Lakonishok Shleifer and Vishny (1994)	1968-1989	-1	1
Sales Growth (3 years)	sale_gr3	Lakonishok Shleifer and Vishny (1994)	1968-1989	-1	1
Sales growth (1 quarter)	saleq_gr1		1967-2016	-1	0
Years 2-5 lagged returns, nonannual	seas_2_5na	Heston and Sadka (2008)	1965-2002	-1	1

Low Leverage

Firm age	age	Jiang Lee and Zhang (2005)	1965-2001	-1	1
----------	-----	----------------------------	-----------	----	---

Liquidity of market assets	aliq_mat	Ortiz-Molina and Phillips (2014)	1984-2006	-1	0
Book leverage	at_be	Fama and French (1992)	1963-1990	-1	0
The high-low bid-ask spread	bidaskhl_21d	Corwin and Schultz (2012)	1927-2006	1	1
Cash-to-assets	cash_at	Palazzo (2012)	1972-2009	1	0
Net debt-to-price	netdebt_me	Penman Richardson and Tuna (2007)	1962-2001	-1	1
Earnings volatility	ni_ivol	Francis et al. (2004)	1975-2001	1	0
R&D-to-sales	rd_sale	Chan Lakonishok and Sougiannis (2001)	1975-1995	1	0
R&D capital-to-book assets	rd5_at	Li (2011)	1952-2004	1	0
Asset tangibility	tangibility	Hahn and Lee (2009)	1973-2001	1	0
Altman Z-score	z_score	Dichev (1998)	1981-1995	1	1

Low Risk

Market Beta	beta_60m	Fama and MacBeth (1973)	1935-1968	-1	1
Dimson beta	beta_dimson_21d	Dimson (1979)	1955-1974	-1	0
Frazzini-Pedersen market beta	betabab_1260d	Frazzini and Pedersen (2014)	1926-2012	-1	1
Downside beta	betadown_252d	Ang Chen and Xing (2006)	1963-2001	-1	1
Earnings variability	earnings_variability	Francis et al. (2004)	1975-2001	-1	0
Idiosyncratic volatility from the CAPM (21 days)	ivol_capm_21d		1967-2016	-1	0
Idiosyncratic volatility from the CAPM (252 days)	ivol_capm_252d	Ali Hwang and Trombley (2003)	1976-1997	-1	1
Idiosyncratic volatility from the Fama-French 3-factor model	ivol_ff3_21d	Ang et al. (2006)	1963-2000	-1	1
Idiosyncratic volatility from the q-factor model	ivol_hxz4_21d		1967-2016	-1	0
Cash flow volatility	ocfq_saleq_std	Huang (2009)	1980-2004	-1	1
Maximum daily return	rmax1_21d	Bali Cakici and Whitelaw (2011)	1962-2005	-1	1
Highest 5 days of return	rmax5_21d	Bali, Brown, Murray and Tang (2017)	1993-2012	-1	1
Return volatility	rvol_21d	Ang et al. (2006)	1963-2000	-1	1
Years 6-10 lagged returns, nonannual	seas_6_10na	Heston and Sadka (2008)	1965-2002	-1	1
Share turnover	turnover_126d	Datar Naik and Radcliffe (1998)	1963-1991	-1	1
Number of zero trades with turnover as tiebreaker (1 month)	zero_trades_21d	Liu (2006)	1963-2003	1	0
Number of zero trades with turnover as tiebreaker (6 months)	zero_trades_126d	Liu (2006)	1963-2003	1	1
Number of zero trades with turnover as tiebreaker (12 months)	zero_trades_252d	Liu (2006)	1963-2003	1	1

Momentum

Current price to high price over last year	prc_highprc_252d	George and Hwang (2004)	1963-2001	1	1
--	------------------	-------------------------	-----------	---	---

Residual momentum t-6 to t-1	resff3_6_1	Blitz Huij and Martens (2011)	1930-2009	1	1
Residual momentum t-12 to t-1	resff3_12_1	Blitz Huij and Martens (2011)	1930-2009	1	1
Price momentum t-3 to t-1	ret_3_1	Jegadeesh and Titman (1993)	1965-1989	1	1
Price momentum t-6 to t-1	ret_6_1	Jegadeesh and Titman (1993)	1965-1989	1	1
Price momentum t-9 to t-1	ret_9_1	Jegadeesh and Titman (1993)	1965-1989	1	1
Price momentum t-12 to t-1	ret_12_1	Jegadeesh and Titman (1993)	1965-1989	1	1
Year 1-lagged return, nonannual	seas_1_1na	Heston and Sadka (2008)	1965-2002	1	1

Profit Growth

Change sales minus change Inventory	dsale_dinv	Abarbanell and Bushee (1998)	1974-1988	1	1
Change sales minus change receivables	dsale_drec	Abarbanell and Bushee (1998)	1974-1988	-1	0
Change sales minus change SG&A	dsale_dsga	Abarbanell and Bushee (1998)	1974-1988	1	0
Change in quarterly return on assets	niq_at_chg1		1972-2016	1	0
Change in quarterly return on equity	niq_be_chg1		1967-2016	1	0
Standardized earnings surprise	niq_su	Foster Olsen and Shevlin (1984)	1974-1981	1	1
Change in operating cash flow to assets	ocf_at_chg1	Bouchard, Krueger, Landier and Thesmar (2019)	1990-2015	1	1
Price momentum t-12 to t-7	ret_12_7	Novy-Marx (2012)	1925-2010	1	1
Labor force efficiency	sale_emp_gr1	Abarbanell and Bushee (1998)	1974-1988	1	0
Standardized Revenue surprise	saleq_su	Jegadeesh and Livnat (2006)	1987-2003	1	1
Year 1-lagged return, annual	seas_1_1an	Heston and Sadka (2008)	1965-2002	1	1
Tax expense surprise	tax_gr1a	Thomas and Zhang (2011)	1977-2006	1	1

Profitability

Coefficient of variation for dollar trading volume	dolvol_var_126d	Chordia Subrahmanyam and Anshuman (2001)	1966-1995	-1	1
Return on net operating assets	ebit_bev	Soliman (2008)	1984-2002	1	1
Profit margin	ebit_sale	Soliman (2008)	1984-2002	1	1
Pitroski F-score	f_score	Piotroski (2000)	1976-1996	1	1
Return on equity	ni_be	Haugen and Baker (1996)	1979-1993	1	1
Quarterly return on equity	niq_be	Hou Xue and Zhang (2015)	1972-2012	1	1
Ohlson O-score	o_score	Dichev (1998)	1981-1995	-1	1
Operating cash flow to assets	ocf_at	Bouchard, Krueger, Landier and Thesmar (2019)	1990-2015	1	1
Operating profits-to-book equity	ope_be	Fama and French (2015)	1963-2013	1	1
Operating profits-to-lagged book equity	ope_bell		1967-2016	1	0
Coefficient of variation for share turnover	turnover_var_126d	Chordia Subrahmanyam and Anshuman (2001)	1966-1995	-1	1

Quality

Capital turnover	at_turnover	Haugen and Baker (1996)	1979-1993	1	0
Cash-based operating profits-to-book assets	cop_at		1967-2016	1	0
Cash-based operating profits-to-lagged book assets	cop_atl1	Ball et al. (2016)	1963-2014	1	1
Change gross margin minus change sales	dgp_dsale	Abarbanell and Bushee (1998)	1974-1988	1	0
Gross profits-to-assets	gp_at	Novy-Marx (2013)	1963-2010	1	1
Gross profits-to-lagged assets	gp_atl1		1967-2016	1	0
Mispricing factor: Performance	mispricing_perf	Stambaugh and Yuan (2016)	1967-2013	1	1
Number of consecutive quarters with earnings increases	ni_inc8q	Barth Elliott and Finn (1999)	1982-1992	1	0
Quarterly return on assets	niq_at	Balakrishnan Bartov and Faurel (2010)	1976-2005	1	1
Operating profits-to-book assets	op_at		1963-2013	1	1
Operating profits-to-lagged book assets	op_atl1	Ball et al. (2016)	1963-2014	1	1
Operating leverage	opex_at	Novy-Marx (2011)	1963-2008	1	1
Quality minus Junk: Composite	qmj	Assness, Frazzini and Pedersen (2018)	1957-2016	1	1
Quality minus Junk: Growth	qmj_growth	Assness, Frazzini and Pedersen (2018)	1957-2016	1	1
Quality minus Junk: Profitability	qmj_prof	Assness, Frazzini and Pedersen (2018)	1957-2016	1	1
Quality minus Junk: Safety	qmj_safety	Assness, Frazzini and Pedersen (2018)	1957-2016	1	1
Assets turnover	sale_bev	Soliman (2008)	1984-2002	1	1
<u>Seasonality</u>					
Market correlation	corr_1260d	Assness, Frazzini, Gormsen, Pedersen (2020)	1925-2015	-1	1
Coskewness	coskew_21d	Harvey and Siddique (2000)	1963-1993	-1	1
Net debt issuance	dbnetis_at	Bradshaw Richardson and Sloan (2006)	1971-2000	-1	1
Kaplan-Zingales index	kz_index	Lamont Polk and Saa-Requejo (2001)	1968-1995	1	1
Change in long-term investments	lti_gr1a	Richardson et al. (2005)	1962-2001	-1	1
Taxable income-to-book income	pi_nix	Lev and Nissim (2004)	1973-2000	1	1
Years 2-5 lagged returns, annual	seas_2_5an	Heston and Sadka (2008)	1965-2002	1	1
Years 6-10 lagged returns, annual	seas_6_10an	Heston and Sadka (2008)	1965-2002	1	1
Years 11-15 lagged returns, annual	seas_11_15an	Heston and Sadka (2008)	1965-2002	1	1
Years 11-15 lagged returns, nonannual	seas_11_15na	Heston and Sadka (2008)	1965-2002	-1	0
Years 16-20 lagged returns, annual	seas_16_20an	Heston and Sadka (2008)	1965-2002	1	1

Change in short-term investments	sti_gr1a	Richardson et al. (2005)	1962-2001	1	0
		Size			
Amihud Measure	ami_126d	Amihud (2002)	1964-1997	1	1
Dollar trading volume	dolvol_126d	Brennan Chordia and Subrahmanyam (1998)	1966-1995	-1	1
Market Equity	market_equity	Banz (1981)	1926-1975	-1	1
Price per share	prc	Miller and Scholes (1982)	1940-1978	-1	1
R&D-to-market	rd_me	Chan Lakonishok and Sougiannis (2001)	1975-1995	1	1

Short-Term Reversal

Idiosyncratic skewness from the CAPM	iskew_capm_21d		1967-2016	-1	0
Idiosyncratic skewness from the Fama-French 3-factor model	iskew_ff3_21d	Bali Engle and Murray (2016)	1925-2021	-1	1
Idiosyncratic skewness from the q-factor model	iskew_hxz4_21d		1967-2016	-1	0
Short-term reversal	ret_1_0	Jegadeesh (1990)	1929-1982	-1	1
Highest 5 days of return scaled by volatility	rmax5_rvol_21d	Assness, Frazzini, Gormsen, Pedersen (2020)	1925-2015	-1	1
Total skewness	rskew_21d	Bali Engle and Murray (2016)	1925-2021	-1	1

Value

Assets-to-market	at_me	Fama and French (1992)	1963-1990	1	0
Book-to-market equity	be_me	Rosenberg Reid and Lanstein (1985)	1973-1984	1	1
Book-to-market enterprise value	bev_mev	Penman Richardson and Tuna (2007)	1962-2001	1	1
Net stock issues	chcsho_12m	Pontiff and Woodgate (2008)	1970-2003	-1	1
Debt-to-market	debt_me	Bhandari (1988)	1948-1979	1	1
Dividend yield	div12m_me	Litzenberger and Ramaswamy (1979)	1940-1980	1	1
Ebitda-to-market enterprise value	ebitda_mev	Loughran and Wellman (2011)	1963-2009	1	1
Equity duration	eq_dur	Dechow Sloan and Soliman (2004)	1962-1998	-1	1
Net equity issuance	eqnetis_at	Bradshaw Richardson and Sloan (2006)	1971-2000	-1	1
Equity net payout	eqnpo_12m	Daniel and Titman (2006)	1968-2003	1	1
Net payout yield	eqnpo_me	Boudoukh et al. (2007)	1984-2003	1	1
Payout yield	eqpo_me	Boudoukh et al. (2007)	1984-2003	1	1
Free cash flow-to-price	fcf_me	Lakonishok Shleifer and Vishny (1994)	1963-1990	1	1
Intrinsic value-to-market	ival_me	Frankel and Lee (1998)	1975-1993	1	0
Net total issuance	netis_at	Bradshaw Richardson and Sloan (2006)	1971-2000	-1	1
Earnings-to-price	ni_me	Basu (1983)	1963-1979	1	1

Operating cash flow-to-market	ocf_me	Desai Rajgopal and Venkatachalam (2004)	1973-1997	1	1
Sales-to-market	sale_me	Barbee Mukherji and Raines (1996)	1979-1991	1	1

Note: This table shows cluster names as underlined section headings and, for each cluster, a description of the factors included, the variable name used in the code, the original reference, the sample period used in the original reference, the sign of the factor (“1” means “long”, “-1” means “short”), and whether the original reference found the factor to be significant (“1” means “yes”, “0” means “no”). For example, the first value factor “at_me” goes long stocks with high values of assets-to-market and shorts those with low values (and would be done the reverse if the sign was “-1” instead of “1”).

Table AIV. Alpha Across Regions

	Factor	US			Developed ex. US			Emerging		
		α_{OLS}	α_{EB}	$\Pr(\alpha_{EB} < 0)$	α_{OLS}	α_{EB}	$\Pr(\alpha_{EB} < 0)$	α_{OLS}	α_{EB}	$\Pr(\alpha_{EB} < 0)$
1	aliq_mat*	-0.35	-0.31	1.00	-0.33	-0.27	1.00	-0.31	-0.28	1.00
2	dsale_drec*	-0.25	-0.22	1.00	-0.05	-0.11	0.86	-0.23	-0.18	0.95
3	bidaskhl_21d	-0.24	-0.25	1.00	-0.62	-0.42	1.00	-0.50	-0.38	1.00
4	ni_livol*	-0.23	-0.18	0.99	-0.20	-0.13	0.90	0.03	-0.07	0.75
5	at_be*	-0.16	-0.10	0.92	0.06	0.04	0.34	0.12	0.04	0.37
6	age	-0.15	-0.18	1.00	-0.64	-0.44	1.00	-0.59	-0.41	1.00
7	kz_index	-0.13	-0.12	0.93	-0.08	-0.08	0.79	-0.32	-0.15	0.93
8	turnover_var_126d	-0.13	-0.12	0.96	-0.02	-0.03	0.60	0.20	0.01	0.46
9	prc	-0.11	-0.02	0.60	0.02	0.06	0.26	0.07	0.07	0.23
10	sti_gr1a*	-0.09	-0.02	0.60	-0.01	0.03	0.37	0.17	0.08	0.21
11	dolvol_var_126d	-0.07	-0.07	0.85	-0.07	-0.03	0.63	0.20	0.02	0.43
12	dsale_dsga*	-0.07	-0.02	0.58	0.20	0.13	0.11	0.33	0.15	0.09
13	ni_ar1*	-0.02	-0.07	0.81	0.04	0.03	0.39	-0.29	-0.06	0.71
14	sale_emp_gr1*	-0.01	-0.03	0.64	-0.24	-0.10	0.83	0.11	-0.00	0.51
15	netdebt_me	-0.01	0.03	0.32	0.11	0.11	0.13	0.20	0.12	0.13
16	z_score	-0.00	0.03	0.36	-0.02	0.03	0.36	0.20	0.10	0.17
17	iskew_hxz4_21d*	0.01	-0.08	0.80	-0.50	-0.18	0.94	-0.37	-0.16	0.90
18	rd_sale*	0.01	0.06	0.22	0.22	0.18	0.03	0.25	0.15	0.08
19	market_equity	0.02	0.13	0.04	0.13	0.21	0.02	0.55	0.37	0.00
20	cash_at*	0.04	0.07	0.18	0.07	0.10	0.15	0.24	0.14	0.08
21	ami_126d	0.05	0.14	0.03	0.14	0.21	0.02	0.38	0.28	0.00
22	ncol_gr1a*	0.05	-0.01	0.57	-0.07	0.02	0.42	-0.08	0.02	0.41
23	debt_me	0.09	0.06	0.21	0.03	0.05	0.31	-0.06	-0.00	0.51
24	ni_inc8q*	0.10	0.13	0.07	0.38	0.25	0.02	0.26	0.18	0.07
25	tax_gr1a	0.10	0.09	0.13	0.06	0.10	0.17	0.30	0.16	0.06
26	saleq_gr1*	0.11	-0.03	0.66	0.03	0.04	0.37	-0.70	-0.16	0.91
27	ret_60_12	0.12	0.03	0.31	0.11	0.19	0.03	0.37	0.30	0.00
28	rd5_at*	0.12	0.17	0.01	0.35	0.29	0.00	0.63	0.31	0.01
29	coskew_21d	0.12	0.12	0.05	0.28	0.19	0.02	-0.04	0.08	0.20
30	saleq_su	0.12	0.16	0.04	0.47	0.30	0.01	0.54	0.29	0.01
31	col_gr1a	0.13	0.01	0.46	0.00	0.07	0.23	-0.11	0.05	0.31
32	iskew_ff3_21d	0.13	0.08	0.18	-0.20	0.01	0.46	0.15	0.11	0.17
33	tangibility*	0.13	0.17	0.01	0.26	0.25	0.00	0.41	0.28	0.00
34	lti_gr1a	0.15	0.09	0.13	-0.01	0.04	0.35	-0.17	-0.01	0.52
35	pi_nix	0.16	0.14	0.05	0.06	0.11	0.14	0.09	0.12	0.13
36	bev_mev	0.17	0.16	0.02	0.28	0.25	0.00	0.27	0.22	0.02
37	gp_atl1*	0.18	0.20	0.01	0.22	0.24	0.01	0.53	0.31	0.00
38	at_me*	0.18	0.16	0.01	0.20	0.20	0.02	0.23	0.19	0.04
39	seas_16_20na	0.19	0.04	0.32	-0.13	0.02	0.44	-0.32	-0.02	0.57
40	zero_trades_21d*	0.19	0.16	0.01	0.21	0.18	0.04	0.40	0.26	0.01
41	ebit_sale	0.20	0.18	0.01	0.21	0.18	0.03	0.25	0.15	0.07
42	ret_3_1	0.21	0.11	0.06	0.26	0.15	0.06	0.08	0.08	0.21
43	be_me	0.23	0.22	0.00	0.32	0.29	0.00	0.32	0.27	0.01
44	op_atl1	0.23	0.23	0.00	0.25	0.25	0.01	0.43	0.27	0.00

45	at_turnover*	0.23	0.23	0.00	0.23	0.24	0.01	0.37	0.25	0.01
46	opex_at	0.24	0.22	0.00	0.22	0.22	0.01	0.23	0.18	0.04
47	ope_bel1*	0.24	0.23	0.00	0.30	0.27	0.00	0.43	0.26	0.00
48	earnings_variability*	0.24	0.16	0.02	0.09	0.07	0.24	-0.00	0.06	0.28
49	seas_1_1na	0.26	0.17	0.01	0.23	0.14	0.06	0.28	0.21	0.02
50	dolvol_126d	0.26	0.34	0.00	0.27	0.35	0.00	0.53	0.43	0.00
51	be_gr1a	0.26	0.13	0.05	0.09	0.15	0.06	-0.10	0.10	0.16
52	div12m_me	0.27	0.27	0.00	0.51	0.46	0.00	0.71	0.51	0.00
53	sale_me	0.28	0.26	0.00	0.35	0.33	0.00	0.42	0.34	0.00
54	ocfq_saleq_std	0.28	0.25	0.00	0.38	0.30	0.00	0.69	0.39	0.00
55	niq_at	0.29	0.36	0.00	0.75	0.57	0.00	0.96	0.58	0.00
56	sale_gr3	0.29	0.16	0.02	0.09	0.19	0.03	0.15	0.23	0.02
57	sale_gr1	0.30	0.16	0.02	0.17	0.20	0.02	-0.19	0.08	0.22
58	ni_be	0.30	0.28	0.00	0.42	0.33	0.00	0.31	0.23	0.01
59	ivol_capm_252d	0.30	0.26	0.00	0.41	0.30	0.00	0.21	0.23	0.01
60	seas_2_5na	0.31	0.22	0.00	0.36	0.41	0.00	0.67	0.54	0.00
61	ret_6_1	0.31	0.22	0.00	0.33	0.23	0.01	0.38	0.29	0.00
62	seas_11_15na*	0.31	0.22	0.00	-0.07	0.08	0.23	-0.26	0.05	0.34
63	o_score	0.31	0.29	0.00	0.38	0.33	0.00	0.45	0.31	0.00
64	beta_dimson_21d*	0.31	0.26	0.00	0.33	0.24	0.00	0.06	0.14	0.08
65	aliq_at*	0.31	0.17	0.01	0.04	0.14	0.08	-0.07	0.12	0.13
66	dgp_dsale*	0.31	0.23	0.00	-0.11	0.02	0.44	-0.03	0.02	0.43
67	rd_me	0.32	0.39	0.00	0.36	0.39	0.00	0.39	0.39	0.00
68	corr_1260d	0.32	0.29	0.00	0.24	0.27	0.00	0.26	0.26	0.00
69	qmj_safety	0.33	0.34	0.00	0.32	0.36	0.00	0.81	0.51	0.00
70	emp_gr1	0.34	0.22	0.00	0.16	0.28	0.00	0.50	0.41	0.00
71	eq_dur	0.34	0.33	0.00	0.38	0.37	0.00	0.52	0.41	0.00
72	betadown_252d	0.35	0.31	0.00	0.48	0.36	0.00	0.25	0.27	0.00
73	prc_highprc_252d	0.35	0.26	0.00	0.38	0.27	0.00	0.40	0.32	0.00
74	turnover_126d	0.36	0.32	0.00	0.50	0.39	0.00	0.50	0.39	0.00
75	ret_1_0	0.36	0.29	0.00	0.20	0.28	0.00	0.07	0.20	0.03
76	gp_at	0.36	0.37	0.00	0.31	0.36	0.00	0.78	0.49	0.00
77	beta_60m	0.37	0.32	0.00	0.38	0.32	0.00	0.43	0.35	0.00
78	taccruals_at	0.37	0.20	0.01	0.06	0.15	0.07	-0.00	0.12	0.13
79	zero_trades_126d	0.38	0.34	0.00	0.52	0.41	0.00	0.51	0.41	0.00
80	ival_me*	0.38	0.35	0.00	0.38	0.37	0.00	0.40	0.36	0.00
81	seas_2_5an	0.39	0.38	0.00	0.55	0.45	0.00	0.46	0.41	0.00
82	ope_be	0.39	0.36	0.00	0.45	0.39	0.00	0.49	0.36	0.00
83	sale_bev	0.40	0.39	0.00	0.32	0.35	0.00	0.59	0.41	0.00
84	niq_be	0.41	0.43	0.00	0.83	0.61	0.00	0.92	0.59	0.00
85	niq_at_chg1*	0.41	0.38	0.00	0.18	0.32	0.00	0.90	0.48	0.00
86	ret_9_1	0.41	0.32	0.00	0.42	0.32	0.00	0.57	0.44	0.00
87	taccruals_ni	0.42	0.22	0.00	0.04	0.12	0.11	-0.18	0.04	0.34
88	dbnetis_at	0.42	0.37	0.00	0.14	0.28	0.00	0.59	0.41	0.00
89	ebit_bev	0.42	0.39	0.00	0.33	0.35	0.00	0.62	0.41	0.00
90	iskew_capm_21d*	0.42	0.33	0.00	-0.04	0.16	0.05	0.21	0.24	0.01
91	eqpo_me	0.42	0.38	0.00	0.37	0.37	0.00	0.43	0.37	0.00

92	seas_16_20an	0.42	0.38	0.00	0.51	0.38	0.00	0.22	0.33	0.00
93	op_at	0.42	0.40	0.00	0.46	0.42	0.00	0.52	0.40	0.00
94	betabab_1260d	0.42	0.39	0.00	0.61	0.50	0.00	0.70	0.53	0.00
95	seas_1_1an	0.42	0.36	0.00	0.47	0.36	0.00	0.11	0.23	0.01
96	ivol_capm_21d*	0.42	0.38	0.00	0.53	0.43	0.00	0.45	0.40	0.00
97	ni_me	0.43	0.41	0.00	0.41	0.43	0.00	0.76	0.55	0.00
98	seas_11_15an	0.43	0.38	0.00	0.20	0.31	0.00	0.39	0.34	0.00
99	at_gr1	0.43	0.29	0.00	0.08	0.22	0.01	0.24	0.31	0.00
100	zero_trades_252d	0.44	0.40	0.00	0.57	0.45	0.00	0.54	0.45	0.00
101	ivol_hxz4_21d*	0.44	0.39	0.00	0.38	0.37	0.00	0.88	0.52	0.00
102	ebitda_mev	0.44	0.43	0.00	0.46	0.46	0.00	0.70	0.53	0.00
103	eqnpo_12m	0.45	0.44	0.00	0.64	0.58	0.00	0.72	0.58	0.00
104	capx_gr3	0.45	0.31	0.00	0.41	0.39	0.00	-0.02	0.25	0.01
105	niq_su	0.47	0.41	0.00	0.24	0.35	0.00	0.78	0.46	0.00
106	rvol_21d	0.47	0.42	0.00	0.52	0.42	0.00	0.33	0.35	0.00
107	ivol_ff3_21d	0.48	0.43	0.00	0.45	0.41	0.00	0.72	0.52	0.00
108	ocf_me	0.49	0.46	0.00	0.31	0.40	0.00	0.92	0.63	0.00
109	ret_12_7	0.49	0.44	0.00	0.54	0.45	0.00	0.41	0.39	0.00
110	capex_abn	0.49	0.35	0.00	0.27	0.32	0.00	0.06	0.27	0.01
111	niq_be_chg1*	0.50	0.44	0.00	0.20	0.35	0.00	0.87	0.50	0.00
112	chcsho_12m	0.50	0.46	0.00	0.42	0.41	0.00	0.28	0.33	0.00
113	resff3_6_1	0.51	0.39	0.00	0.56	0.43	0.00	0.47	0.41	0.00
114	eqnpo_me	0.51	0.47	0.00	0.35	0.40	0.00	0.62	0.49	0.00
115	coa_gr1a	0.51	0.37	0.00	0.17	0.30	0.00	0.26	0.35	0.00
116	qmj_growth	0.52	0.44	0.00	0.24	0.29	0.00	0.25	0.26	0.01
117	eqnetis_at	0.52	0.49	0.00	0.54	0.52	0.00	0.59	0.51	0.00
118	ret_12_1	0.52	0.42	0.00	0.47	0.37	0.00	0.47	0.41	0.00
119	rskew_21d	0.52	0.42	0.00	0.08	0.26	0.01	0.19	0.27	0.00
120	netis_at	0.53	0.50	0.00	0.53	0.52	0.00	0.68	0.55	0.00
121	seas_6_10na	0.53	0.47	0.00	0.65	0.50	0.00	0.36	0.40	0.00
122	mispricing_perf	0.55	0.52	0.00	0.57	0.51	0.00	0.53	0.46	0.00
123	f_score	0.56	0.51	0.00	0.51	0.47	0.00	0.57	0.45	0.00
124	rmax1_21d	0.56	0.50	0.00	0.59	0.48	0.00	0.38	0.40	0.00
125	ocf_at_chg1	0.56	0.49	0.00	0.48	0.45	0.00	0.46	0.43	0.00
126	seas_6_10an	0.57	0.54	0.00	0.99	0.65	0.00	0.33	0.48	0.00
127	fcf_me	0.58	0.56	0.00	0.50	0.55	0.00	1.05	0.73	0.00
128	rmax5_21d	0.58	0.51	0.00	0.63	0.51	0.00	0.31	0.38	0.00
129	capx_gr2	0.60	0.44	0.00	0.35	0.41	0.00	0.10	0.34	0.00
130	qmj_prof	0.60	0.55	0.00	0.48	0.48	0.00	0.59	0.48	0.00
131	qmj	0.60	0.56	0.00	0.55	0.52	0.00	0.62	0.50	0.00
132	capx_gr1*	0.64	0.47	0.00	0.32	0.41	0.00	0.15	0.37	0.00
133	debt_gr3	0.67	0.50	0.00	0.36	0.43	0.00	0.22	0.39	0.00
134	rmax5_rvol_21d	0.67	0.56	0.00	0.34	0.43	0.00	0.04	0.29	0.00
135	lnoa_gr1a	0.69	0.52	0.00	0.29	0.39	0.00	0.07	0.33	0.00
136	fnl_gr1a	0.71	0.54	0.00	0.25	0.40	0.00	0.45	0.48	0.00
137	ppeinv_gr1a	0.71	0.52	0.00	0.24	0.36	0.00	0.07	0.32	0.00
138	oaccruals_at	0.71	0.54	0.00	0.38	0.48	0.00	0.65	0.57	0.00

139	inv_gr1a	0.72	0.55	0.00	0.25	0.39	0.00	0.29	0.44	0.00
140	nfna_gr1a	0.73	0.56	0.00	0.20	0.40	0.00	0.56	0.53	0.00
141	cop_atll	0.75	0.68	0.00	0.46	0.52	0.00	0.77	0.61	0.00
142	dsale_dinv	0.75	0.57	0.00	0.21	0.30	0.00	0.03	0.25	0.01
143	oaccruals_ni	0.76	0.55	0.00	0.15	0.35	0.00	0.57	0.51	0.00
144	ocf_at	0.77	0.71	0.00	0.72	0.65	0.00	0.68	0.58	0.00
145	mispricing_mgmt	0.78	0.63	0.00	0.58	0.64	0.00	0.66	0.68	0.00
146	ncoa_gr1a	0.80	0.61	0.00	0.33	0.45	0.00	0.29	0.47	0.00
147	inv_gr1	0.80	0.62	0.00	0.47	0.52	0.00	0.06	0.38	0.00
148	resff3_12_1	0.81	0.69	0.00	0.98	0.80	0.00	0.72	0.68	0.00
149	nncoa_gr1a	0.82	0.63	0.00	0.28	0.43	0.00	0.30	0.47	0.00
150	cowc_gr1a	0.85	0.64	0.00	0.40	0.51	0.00	0.57	0.57	0.00
151	noa_at	0.87	0.66	0.00	0.36	0.47	0.00	0.22	0.44	0.00
152	noa_gr1a	0.91	0.75	0.00	0.54	0.64	0.00	0.64	0.70	0.00
153	cop_at*	1.02	0.91	0.00	0.62	0.69	0.00	0.85	0.73	0.00

Note: The table shows monthly alpha in percentages across three different regions. α_{OLS} is the intercept from an OLS regression of the factor return on the regional market return. α_{EB} is the factor-region specific posterior mean found via the empirical Bayes procedure applied jointly to all the factor-region specific factors. $\Pr(\alpha_{EB} < 0)$ is the probability that the alpha is negative based on the posterior distribution from the EB procedure. We count a factor as replicated if this probability is below 2.5%. The residual volatility of all strategies have been scaled to 10% annualized. A “*” indicates that the original paper did not propose the factors as a significant predictor of realized returns.

Table AV. Country Information

	Country	MSCI	Start	Stocks	Mega Stocks	Total Market Cap	Median MC
1	USA	Developed	1926-01-31	4,356	449	4.28e+07	796
2	CHN	Emerging	1991-02-28	4,007	102	1.07e+07	816
3	JPN	Developed	1986-01-31	3,867	88	7.05e+06	176
4	HKG	Developed	1986-01-31	2,352	61	5.33e+06	92
5	GBR	Developed	1985-12-31	1,650	36	3.14e+06	158
6	FRA	Developed	1986-01-31	693	36	2.94e+06	121
7	SAU	Emerging	2000-02-29	199	8	2.61e+06	460
8	DEU	Developed	1986-01-31	659	36	2.58e+06	155
9	IND	Emerging	1988-09-30	3,527	26	2.57e+06	10
10	KOR	Emerging	1986-02-28	2,258	21	2.09e+06	131
11	CAN	Developed	1982-03-31	718	35	2.08e+06	198
12	TWN	Emerging	1988-02-29	1,951	12	1.77e+06	121
13	AUS	Developed	1985-11-30	1,702	19	1.68e+06	51
14	CHE	Developed	1986-01-31	233	22	1.66e+06	863
15	NLD	Developed	1986-01-31	119	17	1.11e+06	939
16	SWE	Developed	1986-01-31	688	16	1.00e+06	82
17	ITA	Developed	1986-01-31	359	11	7.68e+05	121
18	ESP	Developed	1986-01-31	182	12	7.02e+05	239
19	RUS	Emerging	1995-08-31	186	11	6.48e+05	184
20	DNK	Developed	1986-01-31	161	8	5.90e+05	184
21	BRA	Emerging	1988-05-31	220	6	5.90e+05	575
22	SGP	Developed	1986-01-31	562	7	5.31e+05	59
23	THA	Emerging	1986-07-31	762	5	5.30e+05	75
24	IDN	Emerging	1989-01-31	656	6	4.93e+05	82
25	MYS	Emerging	1986-01-31	923	2	4.44e+05	55
26	ZAF	Emerging	1986-01-31	261	3	4.07e+05	137
27	BEL	Developed	1986-01-31	131	4	3.94e+05	431
28	NOR	Developed	1986-01-31	272	3	3.44e+05	230
29	MEX	Emerging	1986-02-28	117	3	3.39e+05	992
30	FIN	Developed	1986-01-31	152	6	3.15e+05	218
31	PHL	Emerging	1986-01-31	248	2	2.66e+05	164
32	TUR	Emerging	1990-03-31	389	1	2.38e+05	119
33	ARE	Emerging	2001-06-30	104	4	2.29e+05	284
34	ISR	Developed	1994-12-31	415	1	2.23e+05	110
35	VNM	Frontier	2006-08-31	713	0	1.80e+05	18
36	POL	Emerging	1993-07-31	706	1	1.78e+05	14
37	CHL	Emerging	1989-01-31	174	0	1.71e+05	221
38	QAT	Emerging	2001-12-31	47	2	1.65e+05	971
39	IRL	Developed	1986-01-31	33	4	1.65e+05	652
40	NZL	Developed	1986-01-31	122	0	1.38e+05	196
41	AUT	Developed	1986-01-31	68	0	1.26e+05	327
42	KWT	Frontier	2001-01-31	163	2	9.84e+04	80
43	PER	Emerging	1990-01-31	103	0	8.95e+04	81

44	COL	Emerging	1989-01-31	44	1	8.81e+04	404
45	PRT	Developed	1986-08-31	44	2	8.55e+04	97
46	MAR	Frontier	1995-09-30	71	0	6.51e+04	183
47	GRC	Emerging	1988-09-30	149	0	5.20e+04	33
48	PAK	Emerging	1992-09-30	411	0	4.97e+04	19
49	ARG	Emerging	1988-09-30	68	1	4.95e+04	113
50	NGA	Frontier	1993-11-30	150	0	4.68e+04	11
51	BGD	Frontier	2002-05-31	324	0	4.40e+04	30
52	EGY	Emerging	1996-12-31	204	0	3.87e+04	46
53	HUN	Emerging	1993-06-30	33	0	2.70e+04	54
54	CZE	Emerging	1995-01-31	14	0	2.67e+04	282
55	ROU	Frontier	1997-11-30	76	0	2.55e+04	29
56	BHR	Frontier	2001-03-31	26	0	2.14e+04	243
57	KEN	Frontier	1993-11-30	50	0	2.09e+04	30
58	HRV	Frontier	1997-11-30	73	0	2.07e+04	35
59	BGR	Standalone	1995-12-31	113	0	1.68e+04	24
60	JOR	Frontier	1993-08-31	154	0	1.63e+04	18
61	OMN	Frontier	1998-03-31	107	0	1.62e+04	35
62	LKA	Frontier	1987-06-30	265	0	1.55e+04	15
63	TTO	Standalone	1997-08-31	19	0	1.29e+04	292
64	KAZ	Frontier	2009-06-30	12	0	1.21e+04	394
65	ISL	Standalone	1995-12-31	22	0	1.20e+04	329
66	JAM	Standalone	1993-12-31	66	0	1.13e+04	29
67	SVN	Frontier	1995-03-31	22	0	8.50e+03	122
68	TUN	Frontier	1995-09-30	74	0	8.46e+03	38
69	CIV	Frontier	2002-04-30	39	0	7.61e+03	85
70	MUS	Frontier	1995-08-31	62	0	7.14e+03	48
71	LUX	Not Rated	1986-01-31	8	0	7.14e+03	393
72	LTU	Frontier	1995-10-31	28	0	5.51e+03	61
73	MLT	Standalone	1995-08-31	22	0	5.05e+03	124
74	LBN	Standalone	1997-11-30	8	0	4.12e+03	335
75	EST	Frontier	1996-01-31	19	0	3.48e+03	69
76	TZA	Not Rated	2000-07-31	16	0	3.34e+03	16
77	SRB	Frontier	2009-09-30	29	0	3.17e+03	10
78	BWA	Standalone	1995-09-30	22	0	3.13e+03	91
79	SVK	Not Rated	1986-01-31	10	0	3.12e+03	99
80	CYP	Not Rated	1994-01-31	37	0	3.12e+03	21
81	PSE	Standalone	2008-07-31	27	0	2.97e+03	59
82	GHA	Not Rated	1997-11-30	16	0	2.89e+03	67
83	BMU	Not Rated	2007-08-31	8	0	2.51e+03	118
84	NAM	Not Rated	1996-06-30	8	0	2.18e+03	329
85	MWI	Not Rated	2008-08-31	12	0	2.15e+03	118
86	ECU	Not Rated	1999-04-30	2	0	1.85e+03	927
87	LVA	Not Rated	1997-10-31	20	0	1.17e+03	13
88	UGA	Not Rated	2011-10-31	9	0	1.15e+03	97

89	ZMB	Not Rated	1996-03-31	10	0	5.27e+02	26
90	UKR	Standalone	2008-02-29	4	0	3.28e+02	52
91	GGY	Not Rated	2015-04-30	2	0	2.25e+02	113
92	IRN	Not Rated	2002-05-31	0	0	0.00e+00	0
93	URY	Not Rated	1996-06-30	0	0	0.00e+00	0
All				4.021700e+04	1,092	1.01e+08	

Note: The table shows summary statistics by the country where a security is listed. We include common stocks that are the primary security of the underlying firm, traded on a standard exchange, with non-missing return and market equity data. *Country* is the ISO code of the underlying exchange country. For further information, see https://en.wikipedia.org/wiki/List_of_ISO_3166_country_codes. *MSCI* shows the MSCI classification of each country as of January 7th 2021. For the most recent classification, see <https://www.msci.com/market-classification>. *Start* is the first date with a valid observation. In the next 4 columns, the data is shown as of December 31st 2020. *Stocks* is the number of stocks available. *Mega stocks* is the number of stocks with a market cap above the 80th percentile of NYSE stocks. *Total Market Cap* is the aggregate market cap in million USD. *Median MC* is the median market cap in million USD.

Data Documentation

5.12 Global Factor Data Documentation

We end the Appendix with a documentation of our global factor data and how to use it to replicate our results and for future research. We will continue to update this data and its documentation as seen on our websites. The online document also contains instructions on how to run the code, bug fixes, and so on.

Identifier Variables

Here we define important identifying variables for our empirical analysis. We assign stocks to countries by `excntry`. We assign stocks to size groups via `size_grp`. We only include stocks with 1 on all the the `obs_main`, `exch_main`, `primary_sec` and `common` indicators.

Table AVI. Identifier Variables

Name	Description
<code>excntry</code>	The country of the exchange where the security is traded. Usually expressed as an ISO currency code with the exception of <i>mul</i> which indicates a multi country exchange ⁵⁴
<code>size_grp</code>	This groups each firm into one of five categories: Mega, Large, Small, Micro and Nano cap. The groups are non-overlapping and the breakpoints are based on the market equity of NYSE stocks. In particular, Mega caps are all stocks with market equity larger than the 80th percentile of NYSE stocks, Large caps are all remaining stocks larger than the 50th percentile, Small caps are larger than the 20th percentile, Micro caps are larger than the 1st percentile and Nano caps are the remaining stocks.
<code>obs_main</code>	If there are more than one firm observations for one date, this identifies if the observation is considered as the 'main' observation. If available, CRSP observations are considered as the 'main' observation.
<code>exch_main</code>	Indicator for main exchanges. If CRSP is the source, main exchanges are those with <i>crsp_exchcd</i> 1, 2 and 3. If Compustat is the source, main exchanges are all <i>comp_exchg</i> except 0, 1, 2, 3, 4, 13, 15, 16, 17, 18, 19, 20, 21, 127, 150, 157, 229, 263, 269, 281, 283, 290, 320, 326, 341, 342, 347, 348, 349, 352.
<code>primary_sec</code>	Primary security as identified by Compustat. A 'gvkey' can have up to three different primary securities ('iid') at a given time (US, CA, and international). All observations from CRSP has <code>primary_sec=1</code> .
<code>common</code>	Indicator for common stocks. If CRSP is the source, common is one if the SHRCD variable is 10, 11 or 12. If Compustat is the source, common is one if TPCI is '0'

Helper Functions

This section describes functions that we use to create variables. Many of the functions are used for variables with quarterly, monthly and daily frequencies, and these are specified by “`_zQ`”, “`_zM`” and “`_zD`” respectively, where “`z`” is the number of quarters, months or days that the function is referencing. For example, `COVAR_12M(X, Y)` is the covariance of variables X and Y over the past 12 months.

⁵⁴Typically over the counter exchanges.

Table AVII. Helper Functions

Function	Name	Description
Mean	\bar{X}_z	$\frac{1}{z} \sum_{n=0}^{z-1} X_{t-n}$
Variance	VARC _{.z} (X)	$\frac{1}{z-1} \sum_{n=0}^{z-1} (X_{t-n} - \bar{X}_{tz})^2$
Covariance	COVAR _{.z} (X, Y)	$\frac{1}{z-1} \sum_{n=0}^{z-1} (X_{t-n} - \bar{X}_{tz})(Y_{t-n} - \bar{Y}_{tz})$
Standard Deviation	$\sigma_z(X)$	$\sqrt{VARC_{.z}(X)}$
Skewness	SKEW _{.z} (X)	$\frac{1}{z \times \sigma_z(X)^3} \sum_{n=0}^{z-1} (X_{t-n} - \bar{X}_{tz})^3$
Standardized Unexpected Realization	SUR _{.z} (X)	$\frac{X_t - (X_{t-3} + \frac{(X_{t-3} - X_{t-15})_z}{4})}{\sigma_z(X_{t-3} - X_{t-15})}$
Change to Expectations	CHG.TO_EXP(X)	$\frac{X_t}{(X_{t-12} + X_{t-24})/2}$
Maximum	MAXn _{.z} (X)	The maximum n values of given input.
Quality Minus Junk Helpers		
Earnings Volatility	<i>_EVOL</i>	<i>ROEQ_BE_STD</i> × 2. If this is unavailable, we use <i>ROE_BE_STD</i> .
Rank of Variable	<i>.rVar</i>	Cross-sectional rank of Var within a country ⁵⁵
Z transformation	<i>ZV(rVar)</i>	$\frac{.rVAR - \overline{.rVAR}_z}{.t(.rVAR)}$

Accounting Characteristics

Datasets

- COMP.FUNDA
- COMP.FUNDQ
- COMP.G_FUNDA
- COMP.G_FUNDQ

General Information

- We create characteristics for annual and quarterly accounting data separately. We then take the most recent characteristics value from each dataset to create the final dataset.

⁵⁵*OACCRUALS_AT*, *BETABAB_1260d*, *DEBT_AT* and *_EVOL* are sorted in descending order. All other variables are sorted in ascending order.

- We assume that accounting variables are publically available 4 months after the end of the accounting period.
- In describing accounting variables, we use the Compustat item names from the annual dataset. The equivalent item name in the quarterly dataset can be found by adding a 'q' or 'y' to the end of the annual item name. Specifically, 'q' indicates a value calculated over one quarter while 'y' refers to the cummulative value over the quarters with data available within a fiscal year.

Annualized Accounting Variables from Quarterly Data

- The value of a balance sheet item such as asset or book equity has the same meaning in the annual and the quarterly data. It is the value by the end of a fiscal period.
- The value of an income or cash flow statement item is different. In the annual data, it is calculated over one year. However, in the quarterly data, it is calculated over one quarter. To make quarterly income and cash flows items comparable to the corresponding annual item, we take the sum of the item over the last four quarters.

Accounting Variables

The abbreviation is used to refer to the accounting variable. A suffix of '*' indicates that we have altered the original Compustat item to increase the coverage or to create a variable that is a part of creating a characteristic in the final dataset. The characteristic name will reflect the accounting name except the '*' suffix. As an example, 'gp_at' is gross profit scaled by assets. In general, we will refer to Compustat variables using capital letters.

Table AVIII. Accounting Variables

Name	Abbreviation	Construction
Income Statement		
Sales	sale*	We prefer SALE. If this is unavailable, we use REVT
Cost of Goods Sold	cogs	Compustat item COGS
Gross Profit	gp*	We prefer to use GP. If this is unavailable we use sale*-COGS
Selling, General and Administrative Expenses	xsga	Compustat item XSGA
Advertising Expenses	xad	Compustat item XAD. Note that this is not available in Compustat Global
Research and Development Expenses	xrd	Compustat item XRD. Note that this is not available in Compustat Global
Staff Expenses	xlr	Compustat item XLR
Special Items	spi	Compustat item SPI
Operating Expenses	opex*	We prefer to use XOPR. If this is unavailable, we use COGS+XSGA
Operating Income Before Depreciation	ebitda*	We prefer to use EBITDA. If this is unavailable, we use OIBDP. If this is unavailable, we use SALE*-OPEX*. If this is unavailable, we use GP*-XSGA
Depreciation and Amortization	dp	Compustat Item DP
Operating Income After Depreciation	ebit*	We prefer to use EBIT. If this is unavailable, we use OIADP. If this is unavailable, we use EBITDA*-DP

Name	Abbreviation	Construction
Interest Expenses	int	Compustat item XINT
Operating Profit ala Ball et al (2015)	op*	We use EBITDA* + XRD. If XRD is unavailable, we set it to zero
Operating Profit to Equity	ope*	We use EBITDA*-XINT. Note that we target the same variable as the numerator of the profitability characteristic used to create the Robust-minus weak factor in the fama-French 5 factor model (Fama and French, 2015)
Earnings before Tax and Extraordinary Items	pi*	We prefer to use PI. If this is unavailable we use EBIT*-XINT+SPI+NOPI where we set SPI and NOPI to zero if missing
Income Tax	tax	Compustat item TXT
Extraordinary Items and Discontinued Operations	xido*	We prefer to use XIDO. If this is unavailable, we use XI+DO where we set DO to zero if missing. The reason why we set missing DO to zero is because it is not available in COMP.G_FUNDQ
Net Income	ni*	We prefer to use IB. If this is unavailable, we use NI-XIDO*. If this is unavailable, we prefer PI*-TXT-MII. If MII is unavailable, it is set to zero
Net Income Including Extraordinary Items	nix*	We prefer NI. If this is not available, we prefer NI*+XIDO*. If XIDO* is unavailable, we set it to zero. If that is unavailable, we prefer NI*+XI+DO
Firm Income	fi*	We use NIX*+XINT
Dividends for Common Shareholds	dvc	Compustat Item DVC
Total Dividends	div*	We prefer DVT. If this is not available, we use DV
Income Before Extraordinary Items	ni_qtr*	We use IBQ
Net Sales	sale_qtr*	We use SALEQ
Cash Flow Statement		
Capital Expenditures	capx	Compustat item CAPX
Capital Expenditures to Sales	capex_sale*	We use CAPX / SALE*
Free Cash Flow	fcf*	We use OCF*-CAPX. Note that the free cash flow is computed before financing activities and sale of assets is taken into account
Equity Buyback	eqbb*	We use PRSTKC+PURTSHR Equity Buyback is mainly PRSTKC in NA and PURTSHR in GLOBAL. Either of PRSTKC or PURTSHR are allowed to be missing
Equity Issuance	eqis*	Compustat item SSTK
Equity Net Issuance	eqnetis*	We use EQIS*-EQBB*. Either EQIS* or EQBB* are allowed to be missing
Net Equity Payout	eqpo*	We use DIV*+EQBB*
Equity Net Payout	eqnpo*	We use DIV*-EQNETIS*
Net Long-Term Debt Issuance	dltnetis*	We prefer to use DLTIS-DLTR where we only require that one of the items are non-missing. If this is unavailable, we use LTDCH. If this is unavailable we use the yearly change in long-term book debt DLTT
Net Short-Term Debt Issuance	dstnetis*	We prefer DLCCH. If this is unavailable, we use the yearly change in short-term book debt DLC
Net Debt Issuance	dbnetis*	We use DLTNETIS*+DSTNETIS* and only require one of the items to be non-missing
Net Issuance	netis*	We use EQNETIS*+DBNETIS*. Either EQNETIS* or DBNETIS* are allowed to be missing
Financial Cash Flow	fincf*	We prefer FINCF. If this is unavailable, we use NETIS*-DV+FIAO+TXBCOF. If FIAO or TXBCOF is missing, it is set to zero
Balance Sheet - Assets		

Name	Abbreviation	Construction
Total Assets	at*	We prefer to use AT. If this is unavailable, then we use SEQ* + DLTT + LCT + LO + TXDITC. If LCT, LO, or TXDITC are missing, then they are set to zero
Current Assets	ca*	We prefer ACT. If this is unavailable, we use RECT+INVT+CHE+ACO
Account Receivables	rec	Compustat item RECT
Cash and Short-Term Investment	cash	Compustat item CHE
Inventory	inv	Compustat item INVT
Non-Current Assets	nca*	We use AT* - CA*
Intangible Assets	intan	Compustat item INTAN
Investment and Advances	ivao	Compustat item IVAO
Property, Plans and Equipment Gross	ppeg	Compustat item PPEGT
Property, Plans and Equipment Net	ppen	Compustat item PPENT
Balance Sheet - Liabilities		
Total Liabilities	lt	Compustat item LT
Current Liabilities	cl*	We prefer LCT. If this is unavailable, we use AP+ DLC+ TXP+ LCO
Accounts Payable	ap	Compustat item AP
Short-Term Debt	debtst	Compustat item DLC
Income Tax Payable	txp	Compustat item TXP
Non-Current Liabilities	ncl*	We use LT-CL*
Long-Term Debt	debtlt	Compustat item DLTT
Deferred Taxes and Investment Credit	txditc*	We prefer to use TXDITC. If this is unavailable, we use TXDB+ ITCB
Balance Sheet - Financing		
Preferred Stock	pstk*	We prefer to use PSTKRV. If this is unavailable, we use PSTKL. If this is unavailable, we use PSTK
Total Debt	debt*	We use DLTT+ DLC. Either DLTT or DLC are allowed to be missing
Net Debt	netdebt*	We use DEBT*- CHE where we set CHE to zero if missing
Shareholders Equity	seq*	We prefer to use SEQ. If this is unavailable, we use CEQ+PSTK* where we set PSTK* to zero if missing. If this is unavailable, we use AT- LT
Book Equity	be*	We use SEQ*+TXDITC*-PSTK* where we set TXDITC* and PSTK* to zero if missing
Book Enterprise Value	bev*	We prefer to use ICAPT+DLC-CHE where DLC and CHE are set to zero if missing. If this is unavailable, we use SEQ*+NETDEBT*+ MIB where we set MIB to zero if missing. In the global data ICAPT is reduced by Treasury stock
Balance Sheet - Summary		
Net Working Capital	nwc*	We use CA*-CL*
Current Operating Assets	coa*	We use CA*- CHE
Current Operating Liabilities	col*	We use CL*- DLC. If DLC is missing, it is set to zero
Current Operating Working Capital	cowc*	We use COA*-COL*
Non-Current Operating Assets	ncoa*	We use AT* - CA*- IVAO
Non-Current Operating Liabilities	ncol*	We use LT-CL*- DLTT
Net Non-Current Operating Assets	nmcoa*	We use NCOA*-NCOL*
Financial Assets	fna*	We use IVST+ IVAO. If either is missing, they are set to zero
Financial Liabilities	fnl*	We use DEBT*+PSTK*. If PSTK* is missing, it is set to zero
Net Financial Assets	nfna*	We use FNA*-FNL*
Operating Assets	oa*	We use COA*+NCOA*
Operating Liabilities	ol*	We use COL*+NCOL*
Net Operating Assets	noa*	We use OA*-OL*
Long-Term NOA	lnoa*	PPENT + INTAN + AO - LO + DP

Name	Abbreviation	Construction
Liquid Current Assets	caliq*	We prefer to use CA* - INVT. If this is unavailable, we use CHE + RECT
Property Plant and Equipment Less Inventories	ppeinv*	PPEGT + INVT
Ortiz-Molina and Phillips Liquidity	aliqu*	CHE + 0.75× COA* + 0.5(AT* - CA* - INTAN). If INTAN is missing, we set it to zero
Market Based		
Market Equity	me	We use the market equity for the stock we deem to the primary security of the firm. Importantly, we do not align the market value with the end of the fiscal period. Instead, we update the market value on a monthly basis and align it with the most recently available accounting characteristic
Market Enterprise Value	mev*	We use ME.COMPANY + NETDEBT* × FX*
Market Assets	mat*	We use AT* × FX + BE* × FX + ME.COMPANY
Accruals		
Operating Accruals	oacc*	We prefer NI*-OANCF. If that is unavailable, we use the yearly change in COWC*+the yearly change in NNCOA*
Total Accruals	tacc*	We use OACC* + the yearly change in NFNA*
Operating Cash Flow	ocf*	We prefer to use OANCF. If this is unavailable, we use NI*-OACC*. If this is unavailable, we use NI* + DP - WCAPT. If WCAPT is missing, we use 0.
Quarterly Operating Cash Flow	ocf_qtr*	We use OANCFQ. If this is unavailable, then we use IBQ + DPQ - WCAPTQ. If WCAPTQ is unavailable, we set it to zero
Cash Based Operating Profitability	cop*	We prefer EBITDA*+XRD-OACC*. If XRD is unavailable, we set it to zero
Other		
Employees in Thousands	emp	Compustat item EMP

Table AIX. Accounting Characteristics

Name	Abbreviation	Construction
Growth - Percentage⁵⁶		
Asset Growth 1yr	at_gr1	$\frac{AT^*_t}{AT^*_{t-12}} - 1$
Sales Growth 1yr	sale_gr1	$\frac{SALE^*_t}{SALE^*_{t-12}} - 1$
Sales Growth 3yr	sale_gr3	$\frac{SALE^*_t}{SALE^*_{t-36}} - 1$
Total Debt Growth 3yr	debt_gr3	$\frac{DEBT^*_t}{DEBT^*_{t-36}} - 1$
CAPX 1 year growth	capx_gr1	$\frac{CAPX_t}{CAPX_{t-12}} - 1$
CAPX 2 year growth	capx_gr2	$\frac{CAPX_t}{CAPX_{t-24}} - 1$
CAPX 3 year growth	capx_gr3	$\frac{CAPX_t}{CAPX_{t-36}} - 1$
Quarterly Sales Growth	saleq_gr1	$\frac{SALE_QTR^*_t}{SALE_QTR^*_{t-12}} - 1$

⁵⁶This refers to all variables with a suffix of “_gr1” or “_gr3”. The variables are percentage growth in the accounting variables before the suffix. The number in the suffix refers to either 1 or 3 year growth. For all variables, we only take the percentage growth if the denominator is above zero.

Name	Abbreviation	Construction
Inventory Change 1 yr	inv_gr1	$\frac{INV_t}{INV_{t-12}} - 1$
Sales scaled by Employees Growth 1 yr	sale_emp_gr1	$\frac{SALE_EMP_t}{SALE_EMP_{t-12}} - 1$
Employee Growth 1 yr	emp_gr1	$\frac{EMP_t - EMP_{t-12}}{0.5 \times EMP_t + 0.5 \times EMP_{t-12}}$
Growth - Changed Scaled by Total Assets		
Inventory Change 1yr	inv_gr1a	$\frac{INV_t - INV_{t-12}}{AT^*_t}$
Investment and Advances Change 1yr	lti_gr1a	$\frac{LTI_t - LTI_{t-12}}{AT^*_t}$
Current Operating Assets Change 1yr	coa_gr1a	$\frac{COA^*_t - COA^*_{t-12}}{AT^*_t}$
Current Operating Liabilities Change 1yr	col_gr1a	$\frac{COL^*_t - COL^*_{t-12}}{AT^*_t}$
Non-Current Operating Assets Change 1yr	ncoa_gr1a	$\frac{NCOA^*_t - NCOA^*_{t-12}}{AT^*_t}$
Non-Current Operating Liabilities Change 1yr	ncol_gr1a	$\frac{NCOL^*_t - NCOL^*_{t-12}}{AT^*_t}$
Net Non-Current Operating Assets Change 1yr	nncoa_gr1a	$\frac{NNCOA^*_t - NNCOA^*_{t-12}}{AT^*_t}$
Net Operating Assets Change 1yr	noa_gr1a	$\frac{NOA^*_t - NOA^*_{t-12}}{AT^*_t}$
Financial Liabilities Change 1yr	fnl_gr1a	$\frac{FNL^*_t - FNL^*_{t-12}}{AT^*_t}$
Net Financial Assets Change 1yr	nfna_gr1a	$\frac{NFNA^*_t - NFNA^*_{t-12}}{AT^*_t}$
Effective Tax Rate Change 1yr	tax_gr1a	$\frac{TAX_t - TAX_{t-12}}{AT^*_t}$
Change in Property, Plant and Equipment Less Inventories scaled by lagged Assets	ppeinv_gr1a	$\frac{PPEINV^*_t - PPEINV^*_{t-12}}{AT^*_{t-12}}$
Change in Long-Term NOA scaled by average Assets	lnoa_gr1a	$\frac{LNOA^*_t - LNOA^*_{t-12}}{AT^*_t - AT^*_{t-12}}$
Book Equity Change 1 yr scaled by Assets	be_gr1a	$\frac{BE^*_t - BE^*_{t-12}}{AT^*_t}$
Change in Short-Term Investments scaled by Assets	sti_gr1a	$\frac{IVST_t - IVST_{t-12}}{AT^*_t}$
Profit Margins		
Operating Profit Margin after Depreciation	ebit_sale	$\frac{EBIT^*_t}{SALE^*_t}$
Return on Assets		

Name	Abbreviation	Construction
Gross Profit scaled by Assets	gp_at	$\frac{GP^*_t}{AT^*_t}$
Cash Based Operating Profitability scaled by Assets	cop_at	$\frac{COP^*_t}{AT^*_t}$
Quarterly Income scaled by AT	niq_at	$\frac{NI_{QTR}^*_t}{AT^*_{t-3}}$
Operating Cash Flow scaled by Assets	ocf_at	$\frac{OCF^*_t}{AT^*_t}$
Ball Operating Profit to Assets	op_at	$\frac{OP^*_t}{AT^*_t}$
Ball Operating Profit scaled by lagged Assets	op_at11	$\frac{OP^*_t}{AT^*_{t-12}}$
Gross Profit scaled by lagged Assets	gp_at11	$\frac{GP^*_t}{AT^*_{t-12}}$
Cash Based Operating Profitability scaled by lagged Assets	cop_at11	$\frac{COP^*_t}{AT^*_{t-12}}$
Return on Book Equity		
Operating Profit to Equity scaled by BE	ope_be	$\frac{OPE^*_t}{BE^*_t}$
Net Income scaled by BE	ni_be	$\frac{NI^*_t}{BE^*_t}$
Quarterly Income scaled by BE	niq_be	$\frac{NI_{QTR}^*_t}{BE^*_{t-3}}$
Operating Profit scaled by lagged Book Equity	ope_bell	$\frac{OPE^*_t}{BE^*_{t-12}}$
Return on Invested Capital		
Operating Profit after Depreciation scaled by BEV	ebit_bev	$\frac{EBIT^*_t}{BEV^*_t}$
Issuance		
Net Issuance scaled by Assets	netis_at	$\frac{NETIS^*_t}{AT^*_t}$
Equity Net Issuance scaled by Assets	eqnetis_at	$\frac{EQNETIS^*_t}{AT^*_t}$
Net Debt Issuance scaled by Assets	dbnetis_at	$\frac{DBNETIS^*_t}{AT^*_t}$
Accruals		
Operating Accruals	oaccruals_at	$\frac{OACC^*_t}{AT^*_t}$
Percent Operating Accruals	oaccruals_ni	$\frac{OACC^*_t}{ NIX^*_t }$
Total Accruals	taccruals_at	$\frac{TACC^*_t}{AT^*_t}$
Percent Total Accruals	taccruals_ni	$\frac{TACC^*_t}{ NIX^*_t }$

Name	Abbreviation	Construction
Net Operating Asset to Total Assets	noa_at	$\frac{NOA_t^*}{AT_t^*}$
Financial Soundness Ratios		
Operating Leverage	opex_at	$\frac{OPEX_t^*}{AT_t^*}$
Activity/Efficiency Ratios		
Asset Turnover	at_turnover	$\frac{SALE_t^*}{(AT_t^* + AT_{t-12}^*)/2}$
Miscellaneous		
Sales scaled by BEV	sale_bev	$\frac{SALE_t^*}{BEV_t^*}$
R&D scaled by Sales	rd_sale	$\frac{XRD_t}{SALE_t^*}$
Balance Sheet Fundamental to Market Equity		
Book Equity scaled by Market Equity	be_me	$\frac{BE_t^*}{ME_t}$
Total Assets scaled by Market Equity	at_me	$\frac{AT_t^*}{ME_t}$
Total Debt scaled by ME	debt_me	$\frac{DEBT_t^*}{ME_t}$
Net Debt scaled by ME	netdebt_me	$\frac{NETDEBT_t^*}{ME_t}$
Income Fundamentals to Market Equity		
Net Income scaled by ME	ni_me	$\frac{NI_t^*}{ME_t}$
Sales scaled by ME	sale_me	$\frac{SALE_t^*}{ME_t}$
Operating Cash Flow scaled by ME	ocf_me	$\frac{OCF_t^*}{ME_t}$
Free Cash Flow scaled by ME	fcf_me	$\frac{FCF_t^*}{ME_t}$
R&D scaled by ME	rd_me	$\frac{XRD_t}{ME_t}$
Balance Sheet Fundamentals to Market Enterprise Value		
Book Enterprise Value scaled by MEV	bev_mev	$\frac{BEV_t^*}{MEV_t^*}$
Equity Payout/Issuance to Market Equity		
Net Equity Payout scaled by ME	eqpo_me	$\frac{EQPO_t^*}{ME_t}$
Equity Net Payout scaled by ME	eqnpo_me	$\frac{EQNPO_t^*}{ME_t}$
Income Fundamentals to Market Enterprise Value		
Operating Profit before Depreciation scaled by MEV	ebitda_mev	$\frac{EBITDA_t^*}{MEV_t^*}$

Name	Abbreviation	Construction
Income Growth		
Number of Consecutive Earnings Increases	ni_inc8q	Count number of earnings increases over past 8 quarters
Operating Cash Flow to Assets 1 yr Change	ocf_at_chg1	$OCF_AT_t - OCF_AT_{t-12}$
Change in Quarterly Income scaled by BE	niq_be_chg1	$NIQ_BE_t - NIQ_BE_{t-12}$
Change in Quarterly Income scaled by AT	niq_at_chg1	$NIQ_AT_t - NIQ_AT_{t-12}$
Change Sales minus Change Inventory	dsale_dinv	$CHG_TO_EXP(SALE^*_t) - CHG_TO_EXP(INV_t)$
Change Sales minus Change Receivables	dsale_drec	$CHG_TO_EXP(SALE^*_t) - CHG_TO_EXP(REC_t)$
Change Gross Profit minus Change Sales	dgp_dsale	$CHG_TO_EXP(GP^*_t) - CHG_TO_EXP(SALE^*_t)$
Change Sales minus Change SG&A	dsale_dsga	$CHG_TO_EXP(SALE^*_t) - CHG_TO_EXP(XSGA_t)$
Earnings Surprise	saleq_su	$SUR(SALE_QTR^*)$
Revenue Surprise	niq_su	$SUR(NI_QTR^*)$
Other Variables		
Cash and Short Term Investments scaled by Assets	cash_at	$\frac{CASH_t}{AT^*_t}$
R&D Capital-to-Assets	rd5_at	$\frac{\sum_{n=0}^4 (1 - .2 \cdot n)(XRD_{t-12 \cdot n})}{AT^*_t}$
Age	age	Age of the firms in months
Abnormal Corporate Investment	capex_abn	$\frac{CAPX_SALE^*_t}{(CAPX_SALE^*_{t-12} + CAPX_SALE^*_{t-24} + CAPX_SALE^*_{t-36})/3} - 1$
Earnings before Tax and Extraordinary Items to Net Income Including Extraordinary Items	pi_nix	$\frac{PI^*_t}{NIX^*_t}$
Book Leverage	at_be	$\frac{AT^*_t}{BE^*_t}$
Operating Cash Flow to Sales Quarterly Volatility	ocfq_saleq_std	$SDEV_16Q \left(\frac{OCF_QTR^*_t}{SALE_QTR^*_t} \right)$
Liquidity scaled by lagged Assets	aliq_at	$\frac{ALIQ^*_t}{AT^*_{t-12}}$
Liquidity scaled by lagged Market Assets	aliq_mat	$\frac{ALIQ^*_t}{MAT^*_{t-12}}$
Tangibility	tangibility	$\frac{CASH_t + 0.715 \bullet REC_t + 0.547 \bullet INV_t + 0.535 \bullet PPEG_t}{AT^*_t}$

Name	Abbreviation	Construction
Equity Duration	eq_dur	Outlined in detail here
Piotroski F-Score	f_score	Outlined in detail here
Ohlson O-Score	o_score	Outlined in detail here
Altman Z-Score	z_score	Outlined in detail here
Kaplan-Zingales Index	kz_index	Outlined in detail here
Intrinsic value	intrinsic_value	Outlined in detail here
Intrinsic value-to-market	ival_me	$\frac{INTRINSIC_VALUE^*_t}{ME_t}$
Earnings Variability	earnings_variability	$\frac{\sigma_{60M}(NI^*_t/AT^*_{t-12})}{\sigma_{60M}(OCF^*_t/AT^*_{t-12})}$
1 yr lagged Net Income to Assets	ni_ar1	$\frac{NI^*_{t-12}}{AT^*_{t-12}}$
Net Income Idiosyncratic Volatility	ni_ivol	Outlined in detail here

Market Based Characteristics

Datasets

- CRSP.MSF
- CRSP.DSF
- COMP.SECD
- COMP.G_SECD
- COMP.SECM
- COMP.SECURITY
- COMP.G_SECURITY

Market Variables

A suffix of '*' indicates that we have altered or renamed the original item.

Table AX. Market Variables

Name	Abbreviation	Construction
CRSP Variables ⁵⁷		

⁵⁷lag is a lag function where lag(x) is the value of x from the previous time period

Name	Abbreviation	Construction
Share Adjustment Factor	adjfct*	We use CFACSHR
Shares	shares*	We use SHROUT/1000 so shares outstanding are in millions.
Price	prc*	We use PRC
Local Price	prc_local*	We use PRC*
Highest Daily Price/Ask	prc_high	We use ASKHI. If PRC* or AKSHI are negative, then PRC_HIGH is set to missing
Lowest Daily Price/Bid	prc_low	We use BIDLO. If PRC* or BIDLO are negative, then PRC_LOW is set to missing
Adjusted Proce	prc_adj*	We use PRC* \times ADJFCT*
Market Equity	me*	We use PRC* \times SHARES* so market equity is quoted in million USD.
Company Market Equity	me_company*	We sum ME* grouped by PERMNO and date
Dollar Volume	dolvol*	We use VOL \times PRC*
Return	RET*	We use RET
Local Return	ret_local*	We use RET
Excess Return	ret_exc*	We use (RET*-T30RET)/21. If T30RET is unavailable, we use RF. If the return is a daily return rather than a monthly return, the RET - T30RET is divided by 1 rather than 21.
Excess Return t+1	ret_exc_lead1m*	Excess return (ret_exc*) in month t+1
Time Since Most Recent Return	ret_lag_dif*	We automatically set this to 1
Cumulative Return	ri*	This is the cumulative return estimated from RET*
Monthly Dividend	div_tot*	We use (RET - RETX) \times lag(PRC*) \times (CFACSHR/lag(CFACSHR))
Compustat Variables		
Share Adjustment Factor	adjfct*	We use AJEXDI
Shares	shares*	We use CSHOC/1000000
Price	prc*	We use PRC_LOCAL* \times FX
Local Price	prc_local*	We use PRCCD
Market Equity	me*	We use PRC* \times SHARES*
Company Market Equity	me_company*	We use ME*
Dollar Volume	dolvol*	We use CSHTRD \times PRC*
Return	RET*	We use RET_LOCAL* \times FX
Excess Return	ret_exc*	We use (RET*-T30RET)/21. If T30RET is unavailable, we use RF. If the return is a daily return rather than a monthly return, the RET - T30RET is divided by 1 rather than 21.
Excess Return t+1	ret_exc_lead1m*	Excess return (ret_exc*) in month t+1
Cumulative Return - Local	ri_local*	We use PRC_LOCAL* \times TRFD/AJEXDI
Local Return	ret_local*	We use RILOCAL*/lag(RILOCAL*) - 1
Time Since Most Recent Return	ret_lag_dif*	We estimate the number of days since the previous return. If the returns are monthly rather than daily, then the time is in months
Cumulative Return	ri*	RILOCAL* \times FX*
Monthly Dividend	div_tot*	We use DIV \times FX*. If DIV is missing, we set it to zero
Cash Dividend	div_cash*	We use DIVD \times FX*. If DIVD is unavailable, we set it to zero
Special Cash Dividend	div_spc*	We use DIVSP \times FX*. If DIVSP is unavailable, we set it to zero
Bid-Ask Average Dummy	bidask*	When PRCSTD = 4 then 1, otherwise 0
Asset Pricing Factors		
Excess Market Return	mktrf*	Country specific market return
High Minus Low	hml*	Country specific factor following Fama and French (1993) and using breakpoints from non-micro cap stocks within the country
Small Minus Big ala Fama-French	smb_ff*	Average of small portfolios minus average of large portfolios from hml*

Name	Abbreviation	Construction
Return on Equity	roe*	Country specific factor following Hou, Xue and Zhang (2015) and using breakpoints from non-micro cap stocks within the country. We use double sorts on return on equity and size rather than triple sorts with investment, due to the limited number of stocks in some international markets.
Investment	inv*	Country specific factor following Hou, Xue and Zhang (2015) and using breakpoints from non-micro cap stocks within the country. We use double sorts on investment and size rather than triple sorts with return on equity, due to the limited number of stocks in some international markets
Small Minus Big ala Hou et al	smb.hxz*	Average of small portfolios minus average of large portfolios from roe* and inv*
Market Volatility for Each Stock	_mktvol.zd*	$\sigma_{zD}(MKTRF^*_t)$ ⁵⁸

Table AXI. Market Characteristics

Name	Abbreviation	Construction
Market Based Size Measure		
Market Equity	market_equity	ME^*_t
Equity Payout		
Dividend to Price - 12 Months	div12m.me	$\frac{\sum_{n=0}^{11} DIV_TOT^*_{t-n} \times SHARES^*_{t-n}}{ME^*_t}$
Change in Shares - 12 Month	chsho.12m	$\frac{SHARES^*_t \times ADJFCT^*_t}{SHARES^*_{t-12} \times ADJFCT^*_{t-12}} - 1$
Net Equity Payout - 12 Month	eqnpo.12m	$\log\left(\frac{RI^*_t}{RI^*_{t-12}}\right) - \log\left(\frac{ME^*_t}{ME^*_{t-12}}\right)$
Momentum/Reversal		
Short Term Reversal	ret.1.0	$\frac{RI^*_t}{RI^*_{t-1}} - 1$
Momentum 1-3 Months	ret.3.1	$\frac{RI^*_{t-1}}{RI^*_{t-3}} - 1$
Momentum 1-6 Months	ret.6.1	$\frac{RI^*_{t-1}}{RI^*_{t-6}} - 1$
Momentum 1-9 Months	ret.9.1	$\frac{RI^*_{t-1}}{RI^*_{t-9}} - 1$
Momentum 1-12 Months	ret.12.1	$\frac{RI^*_{t-1}}{RI^*_{t-12}} - 1$
Momentum 7-12 Months	ret.12.7	$\frac{RI^*_{t-7}}{RI^*_{t-12}} - 1$
Momentum 12-60 Months	ret.60.12	$\frac{RI^*_{t-12}}{RI^*_{t-60}} - 1$
Seasonality		
1 Year Annual Seasonality	seas.1.1an	Return in month t-12
2 - 5 Year Annual Seasonality	seas.2.5an	Average return over annual lags from year t-2 to t-5

⁵⁸Must have enough non-missing values of stock to be estimated

Name	Abbreviation	Construction
6 - 10 Year Annual Seasonality	seas.6_10an	Average return over annual lags from year t-6 to t-10
11 - 15 Year Annual Seasonality	seas.11_15an	Average return over annual lags from year t-11 to t-15
16 - 20 Year Annual Seasonality	seas.16_20an	Average return over annual lags from year t-16 to t-20)
1 Year Non-Annual Seasonality	seas.1_1na	Average return from month t-1 to t-11
2 - 5 Year Non-Annual Seasonality	seas.2_5na	Average return over non-annual lags from year t-2 to t-5
6 - 10 Year Non-Annual Seasonality	seas.6_10na	Average return over non-annual lags from year t-6 to t-10
11 - 15 Year Non-Annual Seasonality	seas.11_15na	Average return over non-annual lags from year t-11 to t-15
16 - 20 Year Non-Annual Seasonality	seas.16_20na	Average return over non-annual lags from year t-16 to t-20
Combined Accounting and Market Based Characteristics		
Let e_t be defined as described here		
60 Month CAPM Beta	beta_60m	$\frac{COVAR_{60M}(RET^*_t, MKTRF^*_t)}{VAR_{60M}(MKTRF^*_t)}$
Performance Based Mispricing	mispricing_perf ⁵⁹	$\frac{1}{4}(O_SCORE_t^{r01} + RET_{.12.1}_t^{r01} + GP_AT_t^{r01} + NIQ_AT_t^{r01})$
Management Based Mispricing	mispricing_mgmt	$\frac{1}{6}(CHCSHO_{.12M}_t^{r01} + EQNPO_{.12M}_t^{r01} + OACCRUALS_AT_t^{r01} + NOA_AT_t^{r01} + AT_GR1_t^{r01} + PPEINV_GR1A_t^{r01})$
Residual Momentum - 6 Month	resff3.6.1	$-1 + \prod_{n=1}^6 1 + e_{t-n}$
Residual Momentum - 12 Month	resff3.12.1	$-1 + \prod_{n=1}^{12} 1 + e_{t-n}$
Daily Market Data⁶⁰		
Let ϵ_t be defined as described here		
Return Volatility	rvol_zd	$\sigma_{zD}(RET_EXC^*_t)$
Maximum Return	rmax1_zd	$MAX1_zD(RET^*_t)$
Mean Maximum Return	rmax5_zd	$\frac{1}{5} \sum_{n=1}^5 X_n, X_n \in MAX5_zD(RET^*)$
Return Skewness	rskew_zd	$SKREW_zD(RET_EXC^*_t)$
Price-to-High	prc_highprc_zd	$\frac{PRC_ADJ^*_t}{MAX1_zD(PRC_ADJ^*_t)}$

⁵⁹A rank characteristic has the value of that characteristics rank with respect to other companies' same characteristic of the same month and country scaled [0, 1]. This is identified with a "r01" superscript.

⁶⁰Many of the variables in this section are estimated using rolling windows of data, and the variables are estimated using a variety of window lengths: 21, 126, 252 and 1260 days. In this section, I refer to the number of days as m as a proxy for any of the possible window lengths.

Name	Abbreviation	Construction
Amihud (2002) Measure	ami_zd	$\left(\frac{ RET^*_t }{DOLVOL^*_t} \right)_{zD} * 1000000$
CAPM Beta	beta_zd	Described in detail here
CAPM Idiosyncratic Vol.	ivol_capm_zd	Described in detail here
CAPM Skewness	iskew_capm_zd	Described in detail here
Coskewness	coskew_zd ⁶¹	$\frac{\overline{(\epsilon_t \times MKTRF.DM^2_t)_{zD}}}{\sqrt{(\epsilon_t^2)_{zD} \times (MKTRF.DM^2_t)_{zD}}}$
Fama and French Idiosyncratic Vol.	ivol_ff3_zd	Described in detail here
Fama and French Skewness	iskew_ff3_zd	Described in detail here
Hou, Xue and Zhang Idiosyncratic Vol.	ivol_hxz4_zd	Described in detail here
Hou, Xue and Zhang Skewness	iskew_hxz4_zd	Described in detail here
Dimson Beta	beta_dimson_zd	Created as described in Dimson (1979)
Downside Beta	betadown_zd	Described in detail here
Zero Trades	zero_trades_zd	Number of days with zero trades over period. In case of equal number of zero trading days, turnover_zd will decide on the rank following Liu (2006)
Turnover	turnover_zd	$\left(\frac{TVOL^*_t}{SHARES^*_t * 1000000} \right)_{zD}$
Turnover Volatility	turnover_var_zd	$\frac{\sigma_{zD}((TVOL^*_t / SHARES^*_t) * 1000000)}{TURNOVER.zD_t}$
Dollar Volume	dolvол_zd	$\overline{DOLVOL^*_{t zD}}$
Dollar Volume Volatility	dolvол_var_zd	$\frac{\sigma_{zD}(DOLVOL^*_t)}{DOLVOL.zD_t}$
Correlation to Market	corr_zd	The correlation between $RET_EXC^*_{3l} = RET_EXC^*_t + RET_EXC^*_{t-1} + RET_EXC^*_{t-2}$ and $MKT_EXC_{3l} = MKTRF^*_t + MKTRF^*_{t-1} + MKTRF^*_{t-2}$
Betting Against Beta	betabab_1260d	$\frac{CORR.1260d_t \times RVOL.252d_t}{MKTVOL.252d^*_t}$
Max Return to Volatility	rmax5_rvol_21d	$\frac{RMAX5.21d_t}{RVOL.252d_t}$
21 Day Bid-Ask High-Low	bidaskhl_21d	High-low bid ask estimator created using code from Corwin and Schultz (2012)
Quality Minus Junk		
Quality Minus Junk - Profit	qmj_prof	$ZV(ZV(GP_AT_t) + ZV(NI_BE_t) + ZV(NI_AT_t) + ZV(OCF_AT_t) + ZV(GP_SALE^*_t) + ZV(OACCRUALS_AT_t))$

⁶¹ $MKTRF_DM_t = MKTRF^*_t - MKTRF^*_{t zD}$

Name	Abbreviation	Construction
Quality Minus Junk - Growth	qmj_growth	$ZV\left(ZV(GPOA_CH5_t) + ZV(ROE_CH5_t) + ZV(ROA_CH5_t) + ZV(CFOA_CH5_t) + ZV(GMAR_CH5_t)\right)$
Quality Minus Junk - Safety	qmj_safety	$ZV\left(ZV(BETABAB_1260d_t) + ZV(DEBT_AT_t) + ZV(O_SCORE_t) + ZV(Z_SCORE_t) + ZV(.EVOL_t)\right)$
Quality Minus Junk	qmj	$(QMJ_PROF_t + QMJ_GROWTH_t + QMJ_SAFETY_t)/3$

Detailed Characteristic Construction

This section includes detailed descriptions how we built characteristics that don't easily fit into the Accounting Characteristics or Market Characteristics tables.

- Equity Duration

- Define the following variables:

- * horizon: number of months used to estimate helper variables
- * r: constant used as assumed discount rate
- * roe_mean: constant used as the average ROE value
- * roe_ar1: constant used as the expected growth rate of ROE
- * g_mean: constant used as the average sales growth rate
- * g_ar1: constant used as the expected growth rate of sales

- Create initial variables:

$$\begin{aligned} _roe0 &= \frac{NI^*_t}{BE^*_{t-12}} \\ _g0 &= \frac{SALE^*_t}{SALE^*_{t-12}} - 1 \\ _be0 &= BE^*_t \end{aligned}$$

- * If the number of non-missing observations is less than or equal to 12 or the variables' respective denominators are less than or equal to 1 $_roe0_t$ and $_g0_t$ are set to missing.

- Forecast cash distributions

$$\begin{aligned}
 roe_c &= roe_mean \times (1 - roe_ar1) \\
 g_c &= g_mean \times (1 - g_ar1) \\
 _roe_t &= \sum_{i=1}^{horizon} roe_c + roe_ar1 \times _roe_{t-i} \\
 _g_t &= \sum_{i=1}^{horizon} g_c + g_ar1 \times _g_{t-i} \\
 _be_t &= \sum_{i=1}^{horizon} _be_{t-i} \times (1 + _g_t) \\
 _cd_t &= \sum_{i=1}^{horizon} _be_t \times (_roe_t - _g_t)
 \end{aligned}$$

- Create duration helper variables ⁶²

$$\begin{aligned}
 ed_constant &= horizon + \frac{1+r}{r} \\
 ed_cw_w_t &= \sum_{i=1}^{horizon} ed_cd_w_{i-1} + i \times \frac{_cd_t}{(1+r)^i} \\
 ed_cd_t &= \sum_{i=1}^{horizon} ed_cd_{i-1} + \frac{_cd_t}{(1+r)^i}
 \end{aligned}$$

- Characteristic: $eq_dur_t = \frac{ed_ed_w_t \times FX_t}{ME_COMPANY_t} + ed_constant \times \frac{ME_COMPANY_t - ed_cd_t \times FX_t}{ME_COMPANY_t}$

- Piotroski F-Score

⁶² ed_cw_w , ed_cd and ed_err are equal to 0 at $i = 1$. ed_cw_w and ed_cd recursively build upon themselves over the length of the horizon, so $ed_cw_w_{i-1}$, for example, would be the previous iteration of ed_cw_w

– Create helper variables:

$$\begin{aligned}
-f_roa_t &= \frac{NI^*_t}{AT^*_{t-12}} \\
-f_croa_t &= \frac{OCF^*_t}{AT^*_{t-12}} \\
-f_droa_t &= -f_roa_t - -f_roa_{t-12} \\
-f_acc_t &= -f_croa_t - -f_roa_t \\
-f_lev &= \frac{DLTT_t}{AT^*_t} - \frac{DLTT_{t-12}}{AT^*_{t-12}} \\
-f_liq_t &= \frac{CA^*_t}{CL^*_t} - \frac{CA^*_{t-12}}{CL^*_{t-12}} \\
-f_eqis_t &= EQIS^*_t \\
-f_gm_t &= \frac{GP^*_t}{SALE^*_t} - \frac{GP^*_{t-12}}{SALE^*_{t-12}} \\
-f_aturn_t &= \frac{SALE^*_t}{AT^*_{t-12}} - \frac{SALE^*_{t-12}}{AT^*_{t-24}}
\end{aligned}$$

* For all variables except $-f_acc$, $-f_aturn$ $-f_eqis$, if the count of available observations is less than or equal to 12, then the variable is set to missing. If $-f_aturn$ has less than or equal to 24 non-missing observations, it is set to missing. If a variable has AT^*_t or AT^*_{t-12} as an input and $AT^*_t \leq 0$ or $AT^*_{t-12} \leq 0$, then it is set to missing. If $CL^*_t \leq 0$ or $CL^*_{t-12} \leq 0$ then $-f_liq_t$ is set to missing. If $SALE^*_t \leq 0$ or $SALE^*_{t-12} \leq 0$ then $-f_gm_t$ is set to missing.

– Characteristic⁶³

$$\begin{aligned}
f_score_t &= -f_roa_{>0,t} + -f_croa_{>0,t} + -f_droa_{>0,t} + -f_acc_{>0,t} + \\
&\quad -f_lev_{<0,t} + -f_liq_{>0,t} + -f_eqis_{=0,t} + -f_gm_{>0,t} + -f_aturn_{>0,t}
\end{aligned}$$

• Ohlson O-Score

⁶³A subscript of > 0 , ex: $VAR_{t>0,t}$, is a dummy for if the variable is greater than zero, and it is defined similarly for $VAR_{t<0,t}$ or any other specification. Otherwise, not included as an input, Also, if any variables other than $-f_eqis_t$ are missing, then f_score_t is set to missing.

– Create helper variables:

$$_o_lat_t = AT^*_{t-1}$$

$$_o_lev_t = \frac{DEBT^*_t}{AT^*_t}$$

$$_o_wc_t = \frac{CA^*_t - CL^*_t}{AT^*_t}$$

$$_o_roe_t = \frac{NIX^*_t}{AT^*_t}$$

$$_o_cacl_t = \frac{CL^*_t}{CA^*_t}$$

$$_o_ffo_t = \frac{PI^*_t + DP_t}{LT_t}$$

$$_o_neg_eq_t = 1 \text{ if } LT_t > AT^*_t, \text{ otherwise } 0$$

$$_o_neg_earn_t = 1 \text{ if } NIX^*_t < 0 \text{ and } NIX^*_{t-12} < 0$$

$$_o_nich_t = \frac{NIX^*_t - NIX^*_{t-12}}{|NIX^*_t| + |NIX^*_{t-12}|}$$

- * If $AT^*_t \leq 0$, then $_o_lat_t$, $_o_lev_t$, $_o_wc_t$, and $_o_roe_t$ are set to missing. If $CA^*_t \leq 0$ then $_o_cacl_t$ is set to missing. If $LT_t \leq 0$ then $_o_ffo_t$ is set to missing. If LT_t or AT^*_t are missing, then $_o_neg_eq_t$ is set to missing. If there are less than or equal to 12 observations or either of NIX^*_t and NIX^*_{t-12} are missing, then $_o_nich_t$ and $_o_neg_earn_t$ are set to missing.

– Characteristic:

$$\begin{aligned} o_score_t = & -1.37 - 0.407 \times _o_lat_t + 6.03 \times _o_lev_t + 1.43 \times _o_wc_t + \\ & 0.076 \times _o_cacl_t - 1.72 \times _o_neg_eq_t - 2.37 \times _o_roe_t - \\ & 1.83 \times _o_ffo_t + 0.285 \times _o_neg_earn_t - 0.52 \times _o_nich_t \end{aligned}$$

- Altman Z-Score

- Create helper variables:

$$\begin{aligned} _z_wc_t &= \frac{CA^*_t - CL^*_t}{AT^*_t} \\ _z_re_t &= \frac{RE_t}{AT^*_t} \\ _z_eb_t &= \frac{EBITDA^*_t}{AT^*_t} \\ _z_sa_t &= \frac{SALE^*_t}{AT^*_t} \\ _z_me_t &= \frac{ME_FISCAL_t}{LT_t} \end{aligned}$$

- * If $AT^*_t \leq 0$ then any variable including AT^*_t , then it is set to missing. If $LT_t \leq 0$, then $_z_me_t$ is set to missing.

- Characteristic:

$$_z_score_t = 1.2 \times _z_wc_t + 1.4 \times _z_re_t + 3.3 \times _z_eb_t + 0.6 \times _z_me_t + 1.0 \times _z_sa_t$$

- Kaplan-Zingales Index

- Create helper variables:

$$\begin{aligned} _kz_cf_t &= \frac{NI^*_t + DP_t}{PPENT_{t-12}} \\ _kz_qt &= \frac{AT^*_t + ME_FISCAL_t - BE^*_t}{AT^*_t} \\ _kz_db_t &= \frac{DEBT^*_t}{DEBT^*_t + SEQ^*_t} \\ _kz_dv_t &= \frac{DIV^*_t}{PPENT_{t-12}} \\ _kz_cs_t &= \frac{CHE_t}{PPENT_{t-12}} \end{aligned}$$

- * If the number of non-missing observations is less than or equal to 12, then $_kz_cf_t$, $_kz_dv_t$ and $_kz_cs_t$ are set to zero. If $PPENT_{t-12} \leq 0$ then $_kz_cf_t$, $_kz_dv_t$ and $_kz_cs_t$ are set to missing. If $AT^*_t \leq 0$ then $_kz_qt$ is set to missing. If $(DEBT^*_t + SEQ^*_t) = 0$ then $_kz_db_t$ is set to missing.

- Characteristic:

$$_kz_index = -1.002 \times _kz_cf_t + 0.283 \times _kz_qt + 3.139 \times _kz_db_t - 39.368 \times _kz_dv_t - 1.315 \times _kz_cs_t$$

- Intrinsic Value from Frankel and Lee

- Define r as a constant assumed discount rate
- Create helper variables:

$$\begin{aligned} \textit{iv_po}_t &= \frac{DIV^*_t}{NIX^*_t} \\ \textit{iv_roe}_t &= \frac{NIX^*_t}{(BE^*_t + BE^*_{t-12})/2} \\ \textit{iv_be1}_t &= (1 + (1 - \textit{iv_po}_t) \times \textit{iv_roe}_t) \times BE^*_t \end{aligned}$$

- * If $NIX^*_t \leq 0$ then

$$\textit{iv_po}_t = \frac{DIV^*_t}{AT^*_t \times 0.06}$$

- * If the number of non-missing observations is less than or equal to 12 or $(BE^*_t + BE^*_{t-12}) \leq 0$ then $\textit{iv_roe}_t$ is set to missing.

- Characteristics:

$$\textit{intrinsic_value}_t = BE^*_t + \frac{\textit{iv_roe}_t - r}{1 + r} \times BE^*_t + \frac{\textit{iv_roe}_t - r}{(1 + r) \times r} \times \textit{iv_be1}_t$$

- * If $\textit{intrinsic_value}_t \leq 0$ then it is set to missing.

- Net Income Idiosyncratic Volatility

- Define the following variable ⁶⁴:

$$\textit{ni_at}_t = \frac{NI^*_t}{AT^*_t}$$

- A rolling regression of the following form is run for each company, with the time series split up into n groups:

$$\textit{ni_at}_t = \beta_0 + \beta_1 \textit{ni_at}_{t-12} + u_t$$

where edf_t = the error degrees of freedom of regression and $rmse_t$ = root mean square error of the regression.

- Characteristic:

$$\textit{ni_ivol}_t = \sqrt{\frac{rmse_t^2 \times edf_t}{edf_t + 1}}$$

- Beta, Idiosyncratic Volatility and Skewness of Asset Pricing Factor Regressions

- This section describes the construction of $\textit{beta_zd}$ for the CAPM model, and

⁶⁴If $AT^*_t \leq 0$, then $\textit{ni_at}_t$ is set to missing

the idiosyncratic volatility and skewness characteristics, which are estimated using three different factor models:

- * CAPM (capm):

$$RET_EXC^*_t = \beta_0 + \beta_1 MKTRF^*_t + \epsilon_t$$

- * Fama-French 3 Factor Model (ff3):

$$RET_EXC^*_t = \beta_0 + \beta_1 MKTRF^*_t + \beta_2 HML^*_t + \beta_3 SMB_FF^*_t + e_t$$

- * Hou, Xue and Zhang 4 Factor Model (hxz4):

$$RET_EXC^*_t = \beta_0 + \beta_1 MKTRF^*_t + \beta_2 SMB_HXZ^*_t + \beta_3 ROE^*_t + \beta_4 INV^*_t + \mu_t$$

- Characteristics ⁶⁵:

$$beta_zd = \beta_1 \text{ from the CAPM model}$$

$$ivol_capm_zd_t = \sigma_{zD}(\epsilon_t)$$

$$ivol_ff3_zd_t = \sigma_{zD}(e_t)$$

$$ivol_hxz4_zd_t = \sigma_{zD}(\mu_t)$$

$$iskew_capm_zd_t = SKEW_zD(\epsilon_t)$$

$$iskew_ff3_zd_t = SKEW_zD(e_t)$$

$$iskew_hxz4_zd_t = SKEW_zD(\sigma_t)$$

- Downside Beta

- Define the following regression model run over z days:

$$RET_EXC^*_t = \beta_0 + \beta_1 MKTRF^*_t + \epsilon_t$$

However, we restrict the data to when $MKTRF^*$ is negative.

- Characteristic:

- * $betadown_zd = \beta_1$

⁶⁵_z indicates over how many days the model is run.

FX Conversion Rate Construction

This section outlines how we create a daily dataset, beginning 01/01/1950 to now, of X currency - USD exchange rate using COMPUSTAT. This is run in the macro *compustat_fx()* in the *project_macros.sas* file.

- We use COMP.EXRT_DLY, which has daily conversion rates from GBP to other currencies 'X'.

- Every day available, we estimate the exchange rate fx_t as

$$fx_t = \frac{USD_{GBP,t}}{X_{GBP,t}}$$

where $X_{GBP,t}$ is the exchange rate of GBP to currency X on day t .

- In case there are gaps in information, we assume the exchange rate of the last observation until a new observation is available.

- fx_t is quoted as $\frac{X_t}{USD_t}$, so to go from X to USD, do $X_t \times fx_t$

Chapter 3

Machine Learning and the Implementable Efficient Frontier

with Bryan Kelly, Semyon Malamud, and Lasse Heje Pedersen.

Abstract

We propose that investment strategies should be evaluated based on their net-of-trading-cost return for each level of risk, which we term the “implementable efficient frontier.” While numerous studies use machine learning return forecasts to generate portfolios, their agnosticism toward trading costs leads to excessive reliance on fleeting small-scale characteristics, resulting in poor net returns. We develop a framework that produces a superior frontier by integrating trading-cost-aware portfolio optimization with machine learning. The superior net-of-cost performance is achieved by learning directly about portfolio weights using an economic objective. Further, our model gives rise to a new measure of “economic feature importance.”

Kelly is at Yale School of Management, AQR Capital Management, and NBER. Malamud is at Swiss Finance Institute, EPFL, and CEPR, and is a consultant to AQR. Pedersen is at AQR Capital Management, Copenhagen Business School, and CEPR. We are grateful for helpful comments from Cliff Asness, Jules van Binsbergen, Darrell Duffie, Marc Eskildsen, Markus Ibert, Leonid Spesivtsev, and seminar participants at Copenhagen Business School and the NBER Big Data and Securities Markets Conference, Fall 2021. AQR Capital Management is a global investment management firm that may or may not apply similar investment techniques or methods of analysis as described herein. The views expressed here are those of the authors and not necessarily those of AQR. Semyon Malamud gratefully acknowledges the financial support of the Swiss Finance Institute and the Swiss National Science Foundation, Grant 100018_192692. All four authors appreciate the financial support of INQUIRE Europe.

This paper studies how security information can be used for portfolio selection in a flexible and realistic setting with transaction costs. The goal is thus both to provide a powerful tool for portfolio choice and to shed new light on which security characteristics are economically important drivers of asset pricing.

The financial machine learning (ML) literature provides a flexible framework to combine several characteristics into a single measure of overall expected returns (e.g. Gu et al., 2020b). The same literature documents the relative “feature importance” of different return prediction characteristics (e.g. Chen et al., 2021). These findings suggest that the prediction success of ML methods is often driven by short-lived characteristics that work well for small and illiquid stocks (e.g. Avramov et al., 2021), suggesting that they might be less critical for the real economy (e.g. Binsbergen and Opp, 2019). The high transaction costs of portfolio strategies based on ML imply that these strategies are difficult to implement in practice and, more broadly, raise questions about the relevance and interpretation of the predictability documented in this literature. Do ML-based expected return estimates merely tell us about mispricings that investors don’t bother to arbitrage away because the costs are too large, the mispricing too fleeting, and the relevant stocks too small to matter? Or, do trading-cost-aware ML-based predictions also work for large stocks over significant periods and in a valuable way for large investors, thus providing information about their preferences?

This paper seeks to generate economically useful predictions. We are interested in deriving ML-driven portfolios that can be realistically implemented by market participants with a substantial fraction of aggregate assets under management, such as large pension funds or other professional asset managers. If a strategy is implementable at scale, then the predictive variables that drive such portfolio demands are informative about the equilibrium discount rates of major companies (Kojen and Yogo (2019)).

While ML with transaction costs is challenging to attack with brute force, we deliver a tractable solution through the help of economic theory. Specifically, we show how to integrate the ML problem into a generalized version of the optimal portfolio selection framework of Gârleanu and Pedersen (2013). The main thrust of our approach is to feed the objective function explicit knowledge of implementability, so it knows to search for perhaps subtle but “usable” predictive patterns while discarding more prominent but costly predictive patterns. We develop an ML method to produce optimal portfolios while considering realistic frictions from transaction costs of the securities it trades. Our solution also gives rise to a new measure of “economic feature importance” that captures which characteristics provide the most investment value; in other words, which characteristics contribute the most to the overall portfolio’s risk-adjusted returns after trading costs.

Our approach generalizes Gârleanu and Pedersen (2013) in three important ways. First, while Gârleanu and Pedersen (2013) assume that expected price changes are linear

functions of a set of signals, we allow expected returns to be a fully general function of the signals, opening the door for flexible non-linear ML. Second, our setting is based on stationary returns, not stationary price changes, solving a vexing problem in the portfolio choice literature and providing a new coherence to empirical analysis over long horizons.¹ Third, while Gârleanu and Pedersen (2013) take the data generating process as given, we integrate the estimation process into the optimization process via ML, showing our method’s practical and empirical power.

To understand the difference between our approach and the typical use of ML in finance, note that the latter takes a two-step approach: First, find a function of characteristics that predicts gross returns, and second, use the resulting forecasts to build portfolios. This typical approach abstracts from transaction costs and turnover, and the resultant investment strategies produce negative returns net of transaction costs.

Our approach builds transaction costs directly into the objective function, thus ensuring that the algorithm learns about usable predictability. One element of usable predictability is that it is relevant for large stocks with low transaction costs. Another essential element is alpha decay, that is, how persistent a predictor is. With transaction costs, you will likely own whatever you buy today for a while because the trading costs encourage you only slowly to enter or exit positions. Naturally, understanding the expected return both now and further into the future is relevant.

Empirically, the optimal ML predictor of near-term returns is indeed different from the optimal ML predictor of returns far into the future. In other words, if f_h is the function that best predicts returns h months into the future, $E_t[r_{i,t+h}] = f_h(s_{it})$, then this function is different across h . Given that the standard ML approach uses only f_1 , we see that it misses the information contained in f_h at other horizons, $h > 1$.

One way to implement our approach is to forecast returns across many time horizons $h = 1, 2, \dots$, then to use all of the predictive functions, f_h , appropriately discounted given risk, risk aversion, and the form of the trading cost function. However, this approach requires a highly complex ML formulation to accommodate all predictive functions simultaneously. Using this approach either leads to serious technical challenges (like massive computing costs) or requires cutting essential corners.

Our preferred approach instead learns directly about portfolio weights instead of ex-

¹Gârleanu and Pedersen (2013) show that the portfolio problem simplifies by looking at numbers of shares and price changes because this sidesteps the issue of portfolio growth that has plagued the literature. The portfolio growth is the issue that, if you put 10% of your wealth in IBM stock today, then you will not have 10% of your wealth in IBM next period before trading – because of the price change of IBM and other stocks. Working with the number of shares sidesteps this issue (the number of shares only changes when you trade). Still, the cost is that profit is equal to shares times price changes, so the model cannot be specified in terms of percentage returns, making empirical analysis difficult. We have found a way to work with empirically relevant units and preserve tractability via an approximately optimal solution.

pected returns.² This simple approach delivers an essentially closed-form solution to the highly complex portfolio problem in a single step!

To evaluate the performance of our method, we propose that portfolio choice methods and ML predictions are assessed based on the net-of-cost investment opportunities they produce. Indeed, a fundamental insight in portfolio choice is that investors can depict their investment opportunity set as all the achievable combinations of risk and expected return, giving rise to the “efficient frontier” depicted in most finance textbooks as the tangency line to the hyperbola generated by risky investments. The textbook frontier is drawn in a frictionless setting that abstracts from trading costs, but real-world investors care about their net return. What does the frontier look like when we take trading costs into account?

Panel A of Figure 1 illustrates frontiers for various methods that we study. The baseline for comparison is the cost-agnostic Markowitz-ML solution and the hyperbola of risky investments – both in *gross* terms– that is, our implementation of the textbook frontier using ML. Specifically, these portfolios use ML to build stock-level expected returns, use the academic analog of Barra to build a covariance matrix, and then form ex-ante efficient portfolios from these two inputs. Figure 1.A plots the portfolios’ realized out-of-sample performance. As seen in the figure, while the Markowitz solution is tangent to the hyperbola in a textbook analysis with known means and variances, the Markowitz solution is not precisely tangent in our out-of-sample analysis. In any event, the Markowitz portfolio performs very well out-of-sample, delivering a Sharpe ratio of roughly 2.0 per annum. But this is in *gross* terms. The portfolio’s turnover is enormous and the textbook frontier is non-implementable in practice.

The other lines in Figure 1 show our concept of an “implementable efficient frontier,” that is, the achievable combinations of risk and expected return, net of trading costs. Focusing first on the Markowitz portfolio, we see that its net-of-cost, implementable frontier immediately dives into negative expected return territory as soon as it moves away from a 100% risk-free allocation, as seen in the bottom curve in Panel A. The shape of the implementable efficient frontier may be surprising: Whereas the textbook frontier is a straight line when increasing the allocation to the risky securities while reducing the risk-free allocation (or applying leverage), the true implementable frontier bends down because larger positions incur larger transaction costs. Said differently, we show that the net-of-cost Sharpe ratio declines along the implementable efficient frontier.

To understand the source of the problem for Markowitz-ML at a deeper level, the

²This idea builds on Brandt et al. (2009) who propose using portfolio weights that are parametric (linear) functions of characteristics in a setting without trading costs. We extend their idea to handle the much more complex portfolio problem with transaction costs and by using ML to learn about the optimal weights.

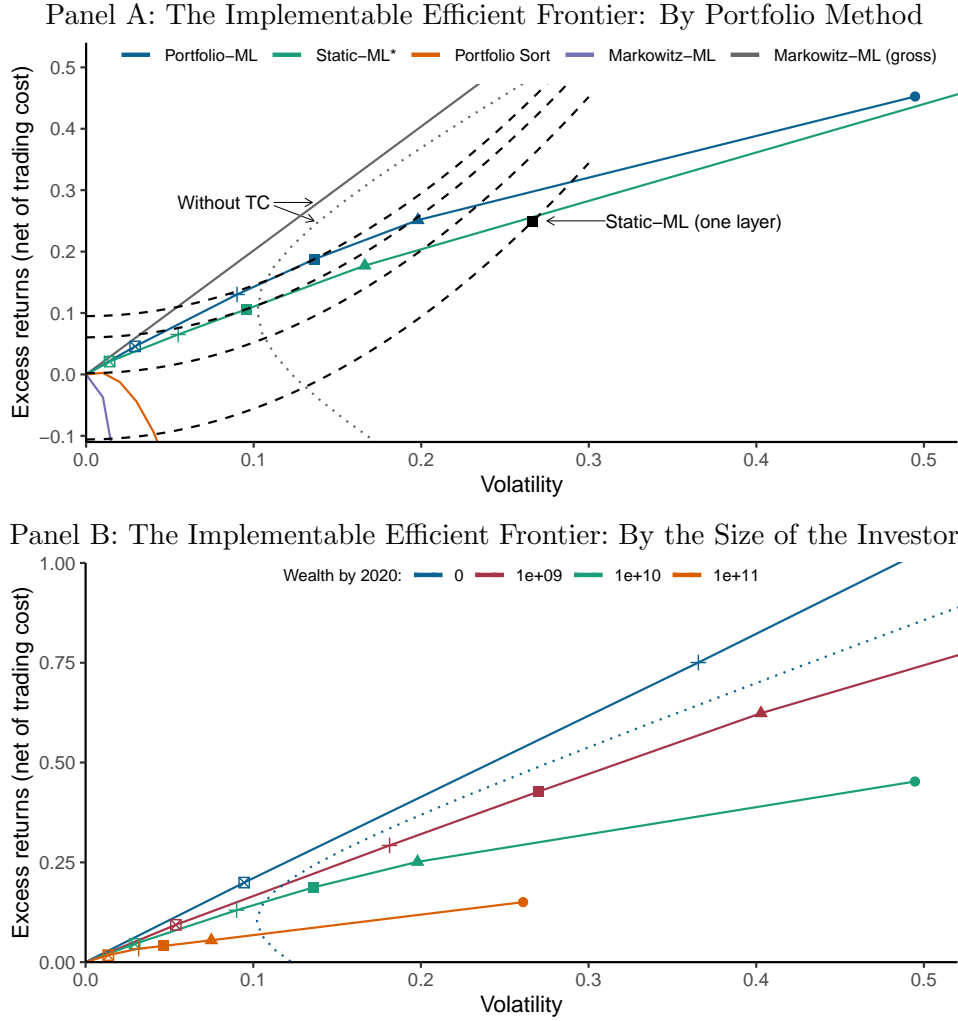


Figure 1. The Implementable Efficient Frontier: Risk vs. Return Net of Trading Costs

Note: Panel A shows the implementable efficient frontier for different portfolio methods with a wealth of \$10B by 2020. The dashed lines show indifference curves. The dotted hyperbola is the mean-variance frontier of risky assets without trading costs, implemented by estimating risk and expected return separately, out-of-sample. The grey line is the Markowitz-ML efficient frontier before trading cost. After trading costs, Markowitz-ML and portfolio sort have downward bending frontiers as these methods are not implementable. Static-ML produces a positive net Sharpe but negative utility, but it works well with an extra tuning layer, denoted Static-ML*. Our Portfolio-ML works significantly better out-of-sample. Panel B shows the implementable efficient frontier at different wealth levels. The dotted hyperbola is the same mean-variance frontier as in Panel A. The blue line is the optimal portfolio of risky and risk-free assets for an investor with zero wealth, corresponding to no trading costs, estimated using our Portfolio-ML method for different relative risk aversions. The blue line would be the tangency line to the hyperbola in a standard in-sample textbook analysis, but it is not exactly tangent out-of-sample. The lower lines illustrate the mean-variance frontiers with larger wealth levels, also estimated using Portfolio-ML. In both panels, the relative risk aversions are 1 (circle), 5 (triangle), 10 (square), 20 (plus), and 100 (boxed cross) and the sample period is 1981-2020. Further details are provided in Section 5.2.

feature importance of this portfolio reveals the culprit: excessive reliance on fleeting small-scale characteristics (e.g., 1-month reversal for small stocks), which bear high turnover, high trading costs, and result in poor net returns. Further, Panel A of Figure 1 also shows

that a standard “portfolio sort” used in the literature is also not implementable.

The difficulty of the standard portfolios from the literature is noteworthy. Still, it is also interesting to compare our approach to a more sophisticated alternative that may be used by some large investors. This sophisticated alternative first uses ML to build expected returns (agnostic of trading costs), then in an additional second-stage optimization, takes transaction costs into account to build portfolios. This “Static-ML” approach delivers a positive net Sharpe but a lower utility than putting all the money in the risk-free asset, as seen from the indifference curves in Figure 1.A.

To create a more difficult benchmark to beat for our preferred method (Portfolio-ML), we further enhance the standard approach by adding several extra hyper-parameters that improve performance by adjusting its scale in various ways. We refer to this approach as Static-ML*, where the “*” indicates that we use an extra tuning stage. Static-ML* performs well, delivering high utility as seen in Figure 1.

Despite that Static-ML* is a sophisticated multi-stage approach that is much more highly parameterized than our Portfolio-ML method, our Portfolio-ML method nevertheless significantly outperforms Static-ML*. To understand why Static-ML* underperforms our approach, note that the first-stage ML procedure produces expected returns dominated by short-term signals. This method does not consider which predictors are persistent and which have quick alpha decay. The second-stage optimization reduces turnover, especially for small stocks, which leads to a much better performance than portfolios that ignore transaction costs. However, this static approach can nevertheless be improved by recognizing the dynamic nature of the portfolio using a method that is sensitive to how expected returns vary across several return horizons (i.e., alpha decay).

In other words, our Portfolio-ML method delivers out-of-sample net-of-cost returns that outperform a highly sophisticated alternative by roughly 20% in Sharpe ratio terms and 60% in utility terms. Further, the feature importance across signals changes when we consider transaction costs. While short-term reversal signals highly influence naive methods, our method seeks to optimally blend return predictability across multiple future horizons, especially for liquid stocks. This leads to value and quality earning the highest feature importance.

Panel B of Figure 1 draws the implementable efficient frontier using our Portfolio-ML at different levels of wealth or asset under management (AUM). Interestingly, while the textbook efficient frontier is the same for all investors, the implementable efficient frontier depends on the investor’s size via the implied trading costs. Indeed, larger investors face worse (i.e., lower) efficient frontiers that “cut into” the hyperbola.

As an interesting benchmark, the top line shows the Portfolio-ML strategy when trading costs are nearly zero since the investor has an AUM near zero. This implementable frontier is obviously good due to the near-zero trading costs, but we note that such so-

phisticated ML-based trading is hardly feasible for small investors in the real world.³

The frontier at each wealth level shows that the set of optimal implementable portfolios is strictly worse for higher AUM investors. This degradation happens for two reasons. First, trading a larger portfolio incurs higher market impact cost. However, the investor can partly mitigate direct transaction costs by trading less, but this increases opportunity costs. Indeed, an investor with larger AUM internalizes price impact from their trades, and this leads the investor to tilt away from highly predictive but costly-to-trade stocks and signals. Large cost-aware investors opt to forego some predictability in order to hold trading costs at bay.

Our paper is related to several large literatures. The first applies machine learning methods to enhance return prediction and enhance portfolio performance, including Gu et al. (2020b), Freyberger et al. (2020), Chen et al. (2021), Kelly et al. (2019), Gu et al. (2021), Jensen et al. (2022b), and Han et al. (2021) in US equity markets; Choi et al. (2021), Leippold et al. (2022), and Cakici and Zaremba (2022) in international equity markets; and Kelly et al. (2022), Bali et al. (2022), and Bali et al. (2021) in bond and derivative markets. Recent empirical papers point out that trading strategies based on these factors and the literature on machine learning in asset pricing cited above involve large transaction costs in practice. This literature includes Li et al. (2020), Chen and Velikov (2021), and Detzel et al. (2021). Motivated by these papers, we develop a flexible portfolio optimization method that lends itself to ML while directly confronting the implementability challenge and explicitly incorporating transaction costs into the ML-based portfolio optimization problem.

The second related literature extends the frictionless paradigm of Markowitz (1952) to study portfolio choice in the presence of transaction costs. Constantinides (1986) and Davis and Norman (1990) analyze settings with a single security, where returns are not predictable. Balduzzi and Lynch (1999) and Lynch and Balduzzi (2000) numerically study single-asset trading with predictable returns and transaction costs. Gârleanu and Pedersen (2013) derive an explicit portfolio solution with multiple assets with predictable returns and transaction costs when returns are driven by a factor model. Gârleanu and Pedersen (2016) extend this to more general dynamics in continuous time and Collin-Dufresne et al. (2020) extend the model to include different liquidity regimes. Our contribution is to derive optimal portfolio rules based on stationary dynamics of returns (rather than dynamic programming with stationary price changes, as in the literature) and fully

³The strategies we develop would be challenging to implement for small investors as they require real-time data on many characteristics across more than a thousand stocks, computation of predictive signals, implementation of ML models, and infrastructure for continually updating and trading these models. Hence, the methods are most relevant for investors large enough to have a staff that can perform these tasks, but, given such capabilities, the implementable investment opportunity set is worse for larger AUM as shown in Figure 1.B.

general functional forms for return predictability while incorporating an arbitrarily large set of predictors.

In summary, we provide a theoretical bridge between portfolio optimization and machine learning with powerful empirical results.

1 Model and the Implementable Efficient Frontier

We consider an economy with N securities traded in discrete time indexed by $t = \dots, -2, -1, 0, 1, 2, \dots$. The return of asset n from time t to $t + 1$ is given by $r_{t+1}^f + r_{n,t+1}$, where r_{t+1}^f is risk-free rate and $r_{n,t+1}$ is the asset's excess return. The vector of all assets' excess returns is denoted r_{t+1} .

An investor observes several characteristics (or signals) for each security, denoted $s_{n,t} \in \mathbb{R}^K$, for example, each asset's valuation ratio, momentum, size, and so on. The characteristics of all assets are collected in the matrix $s_t \in \mathbb{R}^{N \times K}$, and we assume that s_t and r_t are stationary and ergodic. The signals s_t fully characterize the investor's information about returns in a sense that

$$r_{t+1} = \mu(s_t) + \varepsilon_{t+1} \quad (3.1)$$

where the conditional mean $\mu(s_t) = E_t[r_{t+1}]$ and variance $\Sigma(s_t) = \text{Var}_t[r_{t+1}] = \text{Var}_t[\varepsilon_{t+1}]$ are bounded Borel-measurable functions of s_t with Σ being positive definite.

The investor can be seen as a professional asset manager, such as a hedge fund. The investor has wealth or assets under management (AUM) given by w_t at time t . The asset manager's AUM grows at a stochastic rate, g_t^w , so that $w_{t+1} = w_t(1 + g_{t+1}^w)$, which generally depends on performance and on how clients take money in and out of the fund, as specified in Section 1.2. The investor must choose how much capital, $\pi_{n,t}^\$,$ to invest in each asset or, equivalently, choose the fraction of the capital invested in each asset, $\pi_{n,t} = \pi_{n,t}^\$/w_t$. This portfolio choice implies a dollar profit before transaction costs at time $t + 1$ of

$$\text{dollar profit before t-costs}_{t+1} = (r_t^f + r_{t+1})' \pi_t^\$ + (w_t - 1'_N \pi_t^\$) r_t^f = w_t (r_t^f + r'_{t+1} \pi_t) \quad (3.2)$$

where $w_t - 1'_N \pi_t^\$$ is the amount of money in the risk-free money market account, and 1_N is a vector of ones. The corresponding return before trading costs, net of the risk-free rate, is

$$r_{t+1}^{\pi, gross} = \frac{\text{dollar profit before t-costs}_{t+1}}{w_t} - r_t^f = r'_{t+1} \pi_t \quad (3.3)$$

1.1 Trading Costs, Net Returns, and Portfolio Growth

The investor faces transaction costs due to her market impact. Specifically, suppose the investor chooses to trade dollar values given by $\tau_t \in \mathbb{R}^N$ at any time t . This trade leads to a market impact of $\frac{1}{2}\Lambda_t\tau_t$, where $\Lambda_t \in \mathbb{R}^{N \times N}$ is a multivariate version of “Kyle’s Lambda,” which is symmetric and positive semi-definite such that transaction costs are non-negative. It may vary as a function of time and the state s_t of the market.⁴ The resulting transaction cost is the product of the trade size and its market impact, that is,

$$\text{dollar t-costs}_t = \frac{1}{2}\tau_t'\Lambda_t\tau_t. \quad (3.4)$$

To determine the trade size, note that the dollar position $\pi_{n,t-1}^{\$}$ bought at time $t-1$ has grown in value to $\pi_{n,t-1}^{\$}(1+r_t^f+r_{n,t})$. The old dollar position has grown due to the return on the asset (or, said differently, the price change). Hence, the vector of all dollar trade sizes is

$$\begin{aligned} \tau_t &= \pi_t^{\$} - \text{diag}(1+r_t^f+r_t)\pi_{t-1}^{\$} \\ &= w_t\pi_t - w_{t-1}\text{diag}(1+r_t^f+r_t)\pi_{t-1} \\ &= w_t(\pi_t - g_t\pi_{t-1}), \end{aligned} \quad (3.5)$$

where $\text{diag}(v)$ is a diagonal matrix with the vector v in the diagonal and

$$g_t = \text{diag}\left(\frac{1+r_t^f+r_t}{1+g_t^w}\right) \quad (3.6)$$

is the growth of portfolio weights at time t . Combining equations (3.2)–(3.6), we see that the return as a fraction of wealth in excess of the risk-free rate and trading costs is

$$r_{t+1}^{\pi,net} = r_{t+1}^{\pi,gross} - TC_t^{\pi} = r_{t+1}'\pi_t - \frac{w_t}{2}(\pi_t - g_t\pi_{t-1})'\Lambda_t(\pi_t - g_t\pi_{t-1}). \quad (3.7)$$

where $TC_t^{\pi} = \frac{\text{dollar t-costs}_t}{w_t}$. The portfolio’s return naturally depends on the portfolio weights, π , but it also depends on the wealth w_t even though the return is measured in percent of the wealth. This is because trading costs increase by the square of wealth, such that a larger wealth leads to lower portfolio returns after transaction costs. Said differently, a larger investor has a larger market impact (for the same portfolio weights π), thus receiving lower net returns.

⁴The symmetry of Λ is without loss of generality since, if we start with non-symmetric $\tilde{\Lambda}$, we can define $\Lambda = \frac{1}{2}(\tilde{\Lambda} + \tilde{\Lambda}')$ and note that $\tau'\Lambda\tau = \tau'\tilde{\Lambda}\tau$ for any τ .

1.2 Objective Function

The investor maximizes her expected mean-variance utility of portfolio excess returns with relative risk aversion γ_t :

$$\begin{aligned} \text{utility} &= E \left[E_t[r_{t+1}^{\pi, net}] - \text{Var}_t[r_{t+1}^{\pi, net}] \right] \\ &= E \left[\mu(s_t)' \pi_t - \frac{w_t}{2} (\pi_t - g_t \pi_{t-1})' \Lambda_t (\pi_t - g_t \pi_{t-1}) - \frac{\gamma_t}{2} \pi_t' \Sigma (s_t) \pi_t \right]. \end{aligned} \quad (3.8)$$

We make the following assumptions to keep the problem tractable and stationary. First, the investor has constant risk aversion $\gamma_t = \gamma$, and the risk Σ is constant over time. Second, the investor's wealth (or AUM) grows at an exogenous rate (controlled by how clients take money in and out), so the wealth remains a stationary part of the overall market. Specifically, $w_t \Lambda_t = w \Lambda$, such that the investor faces constant transaction costs relative to her wealth. Under these assumptions, the objective function simplifies as follows

$$\text{utility} = E \left[\mu(s_t)' \pi_t - \frac{w}{2} (\pi_t - g_t \pi_{t-1})' \Lambda (\pi_t - g_t \pi_{t-1}) - \frac{\gamma}{2} \pi_t' \Sigma \pi_t \right]. \quad (3.9)$$

where the investor chooses π_t while s_t and g_t are exogenous. In summary, the investor's objective is to maximize utility by choosing her portfolio weights $\pi_t = \pi(\vec{s}_t)$ at each time t as a function of all the signals received up until that time, $\vec{s} = (\dots, s_{t-2}, s_{t-1}, s_t)$.

This setting is ideally suited for a flexible ML implementation for two reasons: First, expected returns are driven by a fully *general function*, μ . Second, the problem is specified in terms of stationary units, namely percentage returns and portfolio weights as fractions of wealth and a stationary objective function.

1.3 The Implementable Efficient Frontier

The utility function depends on risk and expected returns net of trading costs, which gives rise to the implementable efficient frontier as illustrated in Figure 1 in the introduction. Specifically, we defined the implementable efficient frontier as the combination of volatilities and expected net returns, $(\sigma, k(\sigma))_{\sigma \geq 0}$, such that the expected net return is as high as possible for that level of risk:

$$k(\sigma) = \max_{\pi_t \in \Pi} E \left[r_{t+1}^{\pi, net} \right] \quad \text{s.t.} \quad E \left[\pi_t' \Sigma \pi_t \right] = \sigma^2 \quad (3.10)$$

We are mainly interested in the implementable efficient frontier when taking the maximum among all possible portfolios Π , but, as seen in Figure 1, we also consider the frontier among subsets such as standard portfolio sorts.

As an alternative way to compute the implementable efficient frontier, we can derive

the optimal portfolio, π^γ , for any level of risk aversion, γ . Based on all these optimal portfolios, we then compute the corresponding combinations of risk and expected net return:

$$(\sqrt{E[(\pi_t^\gamma)' \Sigma \pi_t^\gamma]}, E[r_{t+1}^{\pi^\gamma, net}])_{\gamma > 0} \quad (3.11)$$

This generates part of the same implementable efficient frontier, as we show in Appendix 7.2. The only difference is that, while (3.10) can generate a downward-sloping curve as seen in Figure 1, (3.11) only produces a part of the frontier that ends before the downward sloping part, since an investor never wants the dominated portfolios on the downward-sloping part. The next result characterizes the frontier.

Proposition 11 (Implementable efficient frontier) *(i) The net Sharpe ratio, $k(\sigma)/\sigma$, is decreasing in σ along the implementable efficient frontier for any level of wealth, $w > 0$, when transaction costs are positive, $\Lambda > 0$; (ii) There exists a critical σ_* such that $k(\sigma)$ is increasing and concave for $\sigma < \sigma_*$; (iii) The part of the frontier $\sigma \in (0, \sigma_*)$ is generated by (3.11) as $\sqrt{E[(\pi_t^\gamma)' \Sigma \pi_t^\gamma]}$ decreases in γ and converges to σ_* when $\gamma \rightarrow 0$; (iv) If $w_1 < w_2$, then the implementable efficient frontier corresponding to a wealth of w_1 is above that of w_2 .*

Interestingly, the implementable efficient frontier has a declining net Sharpe ratio – it is not a straight line with a constant Sharpe ratio as in the textbook frontier without trading costs. The declining net Sharpe ratio reflects that investors cannot freely leverage their portfolio to the desired risk in the presence of trading costs – because more leveraged positions are larger and incur more significant trading costs. Further, larger investors face larger trading costs, leading to a lower frontier. Propositions 12–15 characterize the implementable efficient frontier at a deeper level via the properties of the underlying portfolios.

1.4 Empirical Assessments of Portfolios with Trading Costs

The implementable efficient frontier can be computed in-sample or out-of-sample, where the latter provides a more realistic view of investors' experience, as discussed in our empirical analysis. More broadly, the empirical counterpart of our utility function provides a useful way to evaluate the implementability and economic benefit of any trading strategy:

$$\text{utility}(\text{empirical}) = \frac{1}{T} \sum_{t=1}^T \left[r_{t+1}^{\pi, gross} - TC_t^\pi - \frac{\gamma}{2} \pi_t' \Sigma \pi_t \right]. \quad (3.12)$$

The first part of the sum, $r_t^{\pi, gross} = r'_{t+1}\pi_t$, is simply the average return of the strategy before trading costs. This is the standard metric by which most papers in the literature evaluate trading strategies.

However, real-world trading involves trading costs, so the second term computes the average trading cost over time, $TC_t^\pi = \frac{w}{2} (\pi_t - g_t\pi_{t-1})' \Lambda (\pi_t - g_t\pi_{t-1})$. This term is a bit more complex since it involves both the portfolio weights in the last period, π_{t-1} , and the portfolio in this period, π_t . Specifically, the trading cost is the cost of moving from the grown old portfolio, $g_t\pi_{t-1}$, to the new portfolio. So the first two terms together yield the average return net of trading costs.

Lastly, we need to consider that investors are risk-averse. In particular, two trading strategies that have delivered the same net returns are different if one did so at a much higher risk. Hence, the last term computes the average disutility of risk. Rather than looking at the ex-ante risk, we can also evaluate the ex-post realized risk, $\frac{1}{T} \sum_{t=1} \frac{\gamma}{2} (r_{t+1}^{\pi, net} - \bar{r}^{\pi, net})^2$, where $\bar{r}^{\pi, net}$ is the average net return. Therefore, our utility function suggests that the main object of interest is the average *return net of trading cost and risk*, which can be seen as the utility flow each period:

$$r_{t+1}^{\pi, util} = r_{t+1}^{\pi, gross} - TC_t^\pi - \frac{\gamma}{2} (r_{t+1}^{\pi, net} - \bar{r}^{\pi, net})^2 \quad (3.13)$$

So, when we evaluate trading strategies empirically, we start with each strategy's return gross of costs, $r_{t+1}^{\pi, gross}$, then compute its return net of trading costs, $r_{t+1}^{\pi, net} = r_{t+1}^{\pi, gross} - TC_t^\pi$, and finally focus on the return net of trading costs and risk, $r_{t+1}^{\pi, util}$.

Recall that the net Sharpe ratio declines along the implementable efficient frontier. This result means that an investor cannot just maximize her Sharpe ratio net of trading costs and then choose her risk level – as she could in the standard mean-variance analysis. Instead, she must directly maximize the return net of trading costs and risk, thus jointly considering risk, return, and trading costs. Hence, our framework provides useful tools to evaluate the implementability of trading strategies in general – namely, the concepts of the implementable efficient frontier and the return net of trading cost and risk, $r_{t+1}^{\pi, util}$.

2 Solution

We seek to find the optimal portfolio π_t that maximizes average returns net of trading costs and risk (3.9) or its empirical counterpart (3.12). The problem is too complex and too high-dimensional to attack by brute force ML of a general function $\pi_t = \pi(\vec{s})$ since $\vec{s} = (\dots, s_{t-2}, s_{t-1}, s_t)$ is simply of too high dimension. So, we need help from economic theory before we turn to ML.

2.1 Optimal Dynamic Portfolio Choice

To solve for the optimal portfolio strategy, we use the “discount factor” m defined in the next lemma. To define this discount factor, we use the notation $\bar{g} = E[g_t]$ for the mean portfolio growth rate as defined in (3.6), and $G \in \mathbb{R}^{N \times N}$ for the second moments, $G_{ij} = E[g_{it}g_{jt}]$.

Lemma 3 *The fixed point equation*

$$\tilde{m} = \left(w^{-1} \Lambda^{-1/2} \gamma \Sigma \Lambda^{-1/2} + I + \Lambda^{-1/2} ((\Lambda^{1/2} (I - \tilde{m}) \Lambda^{1/2}) \circ G) \Lambda^{-1/2} \right)^{-1} \quad (3.14)$$

has a unique, symmetric, positive-definite solution $\tilde{m} \in \mathcal{S}(0, 1)$.⁵ For this solution, all eigenvalues of $\Lambda^{-1/2} \tilde{m} \Lambda^{-1/2} \bar{g} \Lambda$ are between zero and one. Furthermore, $m = \Lambda^{-1/2} \tilde{m} \Lambda^{1/2}$ is such that all eigenvalues of $m \Lambda^{-1} \bar{g} \Lambda$ are between zero and one.

We explain in Appendix 7.1 how to calculate m , but for now, let us treat it as a known constant that depends on the exogenous parameters of the model. Based on this known constant, we can compute the optimal portfolio strategy. We start with a simpler case, namely where expected returns are constant. Even in this case, the solution is non-trivial, as is shown by the following proposition.

Proposition 12 (Optimal dynamic strategy: constant expected returns) *Let \tilde{m} be the unique solution to (3.14) in $\mathcal{S}(0, 1)$, and let $m = \Lambda^{-1/2} \tilde{m} \Lambda^{1/2}$. When expected returns, $\mu(s_t) = \bar{\mu} \in \mathbb{R}^N$, as well as g_t^w, r_t^f are constant, then the optimal portfolio is given by*

$$\pi_t = \sum_{\theta=0}^{\infty} \left(\prod_{\tau=1}^{\theta} m g_{t-\tau+1} \right) \frac{1}{w} (I - m \Lambda^{-1} \bar{g} \Lambda)^{-1} m \Lambda^{-1} \bar{\mu} \quad (3.15)$$

Furthermore, it is the unique L_2 -solution to the stochastic difference equation

$$\pi_t = m g_t \pi_{t-1} + \frac{1}{w} (I - m \Lambda^{-1} \bar{g} \Lambda)^{-1} m \Lambda^{-1} \bar{\mu}. \quad (3.16)$$

To understand the intuition for this proposition, note that the optimal portfolio starts with the old grown position, $g_t \pi_{t-1}$, and then trades toward a fixed portfolio. To understand the direction of the trade, it is useful to write the optimal portfolio as

$$\pi_t = m g_t \pi_{t-1} + (I - m) A = g_t \pi_{t-1} + (I - m) (A - g_t \pi_{t-1}), \quad (3.17)$$

⁵The discount factor m is defined by an equation involving the symbol “ \circ ,” which is an element-wise matrix product. The element-wise product is also called the Hadamard product, and, for any two matrices A and B , it is computed as the matrix $(A \circ B)_{i,j} = A_{i,j} B_{i,j}$. Further, $\mathcal{S}(0, 1)$ is the set of positive-definite matrix-valued functions with eigenvalues between zero and one.

where A is the “aim” portfolio. The aim portfolio has the property that, if the investor holds this portfolio, then the investor optimally does not trade, and otherwise the investor trades in the direction of the aim with trading speed $I - m$. We see from Proposition 12 that the aim portfolio is

$$A = (I - m)^{-1}(I - m\Lambda^{-1}\bar{g}\Lambda)^{-1} c \underbrace{\frac{1}{\gamma}\Sigma^{-1}\bar{\mu}}_{\text{Markowitz}}, \quad (3.18)$$

where the matrix c is given by

$$c = \frac{\gamma}{w} m\Lambda^{-1}\Sigma \quad (3.19)$$

So we see that the aim portfolio is related to the Markowitz portfolio but adjusted to account for transaction costs.⁶

We next present a tractable approximation of the optimal portfolio when the expected returns are not too large. Specifically, we write expected returns as $\mu(s_t) = \epsilon\tilde{\mu}(s_t)$, where $\tilde{\mu}$ is a given function and ϵ is a small number that measures the magnitude of expected returns.

Proposition 13 (Optimal dynamic strategy) *Let \tilde{m} be the unique solution to (3.14) in $\mathcal{S}(0, 1)$ and let $m = \Lambda^{-1/2}\tilde{m}\Lambda^{1/2}$. With expected returns $\mu(s_t) = \epsilon\tilde{\mu}(s_t)$ and $g_t^w = g^w + O(\epsilon)$, $r_t^f = r^f + O(\epsilon)$, the optimal portfolio is*

$$\pi_t = m g_t \pi_{t-1} + (I - m)A_t + O(\epsilon^2), \quad (3.20)$$

where the aim portfolio A_t at time t is

$$A_t = (I - m)^{-1} \sum_{\tau=0}^{\infty} (m\Lambda^{-1}\bar{g}\Lambda)^\tau c E_t \left[\underbrace{\frac{1}{\gamma}\Sigma^{-1}\mu(s_{t+\tau})}_{\text{Markowitz}_{t+\tau}} \right]. \quad (3.21)$$

This key theoretical result of the paper shows how to choose the optimal portfolio in two surprisingly simple and intuitive equations. The first equation (3.20) says that one should always start from the grown position inherited from the last period and then trade toward an aim portfolio.

The second equation (3.21) shows how the aim portfolio depends on the current and future Markowitz portfolios, thus providing an optimal risk-return tradeoff along the path where these stocks are expected to remain in the portfolio while simultaneously econo-

⁶We note that, under certain conditions, there is only a small adjustment in the sense that $(I - m)^{-1}(I - m\Lambda^{-1}\bar{g}\Lambda)^{-1} c$ is close to I . For example, this happens in the limit when G is close to $\mathbf{1}\mathbf{1}'$. As we show in the Appendix, when $G = \mathbf{1}\mathbf{1}'$, we have $\bar{g} = \text{diag}(\mathbf{1})$, and $c = (I - m)^2$.

mizing transaction costs.⁷ Proposition 13 generalizes the Gârleanu and Pedersen (2013) portfolio optimization principle, “aim in front of the target,” when facing trading cost frictions. Unlike Gârleanu and Pedersen (2013), we do not require specific assumptions on return dynamics but, instead, allow a general function $\mu(\cdot)$ that predicts returns.

The proposition also leads to several economically intuitive properties, as shown next.

Proposition 14 (Trading speed) *The eigenvalues of the matrix m are monotone decreasing in G , Σ and γ and increasing in w , in the sense of positive semi-definite order.⁸*

To understand the intuition behind these results, recall from (3.20) that m is the persistence of the optimal portfolio, or, equivalently, $I - m$ is the trading speed toward the aim. At the same time, m also determines how much the aim portfolio weights near-time performance versus long-term returns, as seen in (3.21). So, consider what happens when we move to the right in the implementable efficient frontier in Figure 1 by decreasing risk aversion. This decreasing risk aversion means that m increases, thus reducing trading speed and making the aim more focused on persistent signals. In other words, trading costs increase as the investor takes more risk, but the investor compensates by trading more slowly toward a more stable aim. Likewise, an investor with larger wealth w has a lower trading speed because of more significant market impact costs, providing economic intuition for Figure 1.B.

The next proposition considers limiting portfolios with small are large wealth.

Proposition 15 (Small and large investors) *When wealth approaches zero such that transaction costs become negligible, $w \rightarrow 0$, the discount factor converges as $m \rightarrow 0$ and the optimal portfolio policy converges to the Markowitz portfolio, $\pi_t \rightarrow \frac{1}{\gamma} \Sigma^{-1} \mu_t$.*

When wealth grows large, $w \rightarrow \infty$, the optimal portfolio diminishes, $\pi_t \rightarrow 0$, but the discount factor m and rescaled portfolio, $w\pi_t$, and aim portfolio, wA_t , converge to finite limits if Λ is diagonal and $\bar{g}_i = \frac{1+r^f+\mu_i}{1+g^w} > 1$ for all i .

Naturally, a tiny investor holds a portfolio close to the Markowitz portfolio because of the low market impact costs. The limiting behavior as wealth goes to infinity is less obvious: As wealth grows infinite, the investor ultimately holds almost all wealth in the risk-free asset as trading a meaningful proportion of wealth in illiquid assets becomes too costly. However, this result does not mean that the portfolio in dollar terms is not large. Instead, what happens is that the portfolio, measured in dollar terms, grows toward a finite limit. In other words, a maximum amount of money can be made in the market, and as wealth increases, the investors ultimately hold this “maximum dollar portfolio.”

⁷For the convergence of the series (3.21), it is important that $m\bar{g}$ has all eigenvalues below one in absolute value. As we show in the Appendix, a stronger claim holds, and $\Lambda^{1/2}\bar{g}^{1/2}m\bar{g}^{1/2}\Lambda^{-1/2}$ is a symmetric, positive semi-definite matrix with all eigenvalues between zero and one.

⁸Given two symmetric matrices A, B , we write $A \geq B$ in the sense of positive semi-definite order if $A - B$ is positive semi-definite.

2.2 Implementing the Solution with Machine Learning

From a machine learning perspective, Proposition 13 is a powerful result if we assume that s_t is Markovian. To see the power of this result, note that the proposition transfers the ML problem from looking for a general function π all *current and past* signals to a problem of looking for a function $A(s_t)$ of only the *current* signals. This enormous reduction in dimension vastly simplifies the problem. There are dual ways of using machine learning to implement Proposition 13, which we call “Multi-period-ML” and “Portfolio-ML,” respectively. These approaches are designed to find the dynamic optimal portfolio while being aware of transaction costs in a theoretically consistent manner. We describe each approach in turn.

Multi-period-ML: Machine Learning about Expected Returns across Horizons

The first approach to apply Proposition 13 empirically is to compute the aim portfolio using the expected returns across multiple future periods. To understand this approach, recall first that $\mu_t = E_t(r_{t+1})$ is the short-term expected return, so that $E_t[\mu_{t+\tau}] = E_t[E_{t+\tau}[r_{t+\tau+1}]] = E_t[r_{t+\tau+1}]$ is the current expectation about returns τ periods in the future. Using this identity, the aim portfolio can be written as

$$A(s_t) = (I - m)^{-1} \sum_{\tau=0}^{\infty} (m\bar{g})^\tau c \frac{1}{\gamma} \Sigma^{-1} E_t[r_{t+1+\tau}] \quad (3.22)$$

The aim portfolio depends on expected returns across all future time horizons.

One approach to apply Proposition 13 empirically is first to use standard ML techniques to predict returns, but do this for a range of forecasting horizons, thus producing proxies for $E_t[r_{t+1+\tau}]$ for all τ . Using these forecasts, the aim portfolio is given by (3.22). The resulting portfolio can be computed recursively using (3.20), which we call $\pi^{\text{Multi-period-ML}}$ as it is based on expected returns over multiple periods.

While there exist many ML methods that can be used to forecast returns, we focus on a single method throughout the paper for its unique combination of flexibility and simplicity (the appendix contains robustness analysis). Specifically, we use the random features (RF) method of Rahimi and Recht (2007).⁹ This method is based on the insight that any function can be approximated arbitrarily well by a linear combination of known auxiliary functions. In other words, we can write $E_t[r_{i,t+1+\tau}] = f(s_{i,t})b_\tau$, where f is a known family of functions of the signals and b_τ is a parameter to be estimated (Section 4.2 details our empirical methodology). The RF method is powerful in predicting returns and easily adaptable to our second method, discussed next.

⁹Kelly et al. (2022) analyze it in the context of return prediction and portfolio choice.

Portfolio-ML: Machine Learning directly about Portfolio Weights

The disadvantage of Multiperiod-ML is that we need a return prediction model for all future return horizons, not just one period ahead. An alternative, and our preferred approach, is to learn *directly* about portfolio weights rather than the two-step procedure of first predicting returns and then constructing portfolios. We thus refer to this approach as “Portfolio-ML.”

We use (3.12) directly as our ML objective, where the optimal portfolio π depends on the aim A , and then search for the function A that maximizes this objective. This method uses the insights that (i) we can focus on the aim portfolio A , which only depends on current signals, and (ii) the objective should penalize transaction costs.

To see how this works, note that Proposition 13 shows that the optimal portfolio is a weighted average of the inherited position and the current aim portfolio via (3.20). We can express this result as saying that the current optimal portfolio depends on the current and past aim portfolios and their growth over time:

$$\pi_t = \sum_{\theta=0}^{\infty} \left(\prod_{\tau=1}^{\theta} m g_{t-\tau+1} \right) (I - m)A(s_{t-\theta}). \quad (3.23)$$

So, we can replace π by this expression in the objective function (3.12), which leaves us the task of finding the best aim portfolio, $A(\cdot)$, based on an economic objective function.

In other words, we need to find a general function $A(s_t)$ that maximizes the expected utility. To do this, we again use the ML insight that any function can be approximated arbitrarily well by a linear combination of known auxiliary functions. Specifically, we write $A_t = f(s_t)\beta$, where f is a set of known functions (random features) of the signals and $\beta \in \mathbb{R}^p$ is an unknown vector of parameters. For example, if portfolio weights were linear in the signals, we could take f to be the identity such that $A_t = s_t\beta$. We take f as a set of random features, just like we did to predict returns (detailed in Section 4.2).

So we have boiled the portfolio choice problem down to finding the parameter β , and we next show how to do that in closed form. Plugging $A(s_t) = f(s_t)\beta$ into equation (3.23), we see that the optimal portfolio depends on known elements and the unknown parameter β :

$$\pi_t = \left[\sum_{\theta=0}^{\infty} \left(\prod_{\tau=1}^{\theta} m g_{t-\tau+1} \right) (I - m)f(s_{t-\theta}) \right] \beta \equiv \Pi_t \beta \quad (3.24)$$

where Π_t is defined by the last equation. Using this formulation for π_t in the objective

(3.9), we have

$$\begin{aligned}
\text{utility} &= \frac{1}{T} \sum_t \left[r'_{t+1} \pi_t - \frac{\gamma}{2} \pi'_t \Sigma \pi_t - \frac{w}{2} (\pi_t - g_t \pi_{t-1})' \Lambda (\pi_t - g_t \pi_{t-1}) \right] \\
&= \frac{1}{T} \sum_t \left[r'_{t+1} \Pi_t \beta - \frac{\gamma}{2} \beta' \Pi'_t \Sigma \Pi_t \beta - \frac{w}{2} (\Pi_t \beta - g_t \Pi_{t-1} \beta)' \Lambda (\Pi_t \beta - g_t \Pi_{t-1} \beta) \right] \\
&= \frac{1}{T} \sum_t \left[\underbrace{r'_{t+1} \Pi_t \beta}_{\equiv \tilde{r}'_{t+1}} - \frac{1}{2} \beta' \underbrace{[\gamma \Pi'_t \Sigma \Pi_t + w (\Pi_t - g_t \Pi_{t-1})' \Lambda (\Pi_t - g_t \Pi_{t-1})]}_{\equiv \tilde{\Sigma}_t} \beta \right] \\
&\equiv E_T[\tilde{r}'_{t+1}] \beta - \frac{1}{2} \beta' E_T[\tilde{\Sigma}_t] \beta
\end{aligned} \tag{3.25}$$

So we can maximize utility by maximizing this quadratic equation in the unknown parameter β . To ensure a robust solution, we add ridge penalty $-\lambda \beta' \beta$, yielding the following solution:

Proposition 16 (Portfolio-ML) *The aim portfolio can be estimated as $A(s_t) = f(s_t) \beta_T$, and the corresponding optimal portfolio, $\pi^{\text{Portfolio-ML}}$, is given by (3.24), where*

$$\beta_T = (E_T[\tilde{\Sigma}_t] + \lambda I)^{-1} E_T[\tilde{r}_{t+1}] \tag{3.26}$$

Amazingly, this approach delivers a closed-form solution for the optimal dynamic portfolio in light of transaction costs. To find the optimal portfolio, we compute the two “expectations” on the right-hand side of (3.26) as their sample counterparts seen in (3.25). These sample counterparts depend only on data (r_{t+1}, s_t) , known parameters, and the ridge parameter λ , which is chosen via ML validation as discussed in Section 4.2. With enough random features and enough time, the estimated portfolio in Proposition 16 asymptotically recovers the optimal portfolio, as discussed in more detail in the appendix (Proposition 19).

2.3 Economic Feature Importance

It is important to determine which characteristics are economically important. To address this issue, we consider the value function $V(s)$, the maximum utility for a given set of signals s . We define the importance of any feature n as

$$\iota_n = V(s) - V(s^{-n}) \tag{3.27}$$

where s^{-n} is the set of signals s , except that we drop feature n at all times. In other words, the importance of feature n is the drop in utility when the investor no longer has

access to this information. The following results provide an intuitive characterization of the drivers of feature importance.

Proposition 17 *In the limit when Λ is small, the investor's steady-state optimal utility is $V(s) = v(s) - c(s) + O(\|\Lambda\|^2)$, where*

$$v(s) = \frac{1}{2} E[\text{Markowitz}'_t \Sigma \text{Markowitz}_t], \quad (3.28)$$

is the value function without transaction costs, and c measures the cost of time-variation in the Markowitz portfolio:

$$c(s) = \frac{1}{2} E[(\text{Markowitz}_{t+1} - g_{t+1} \text{Markowitz}_t)' \Lambda (\text{Markowitz}_{t+1} - g_{t+1} \text{Markowitz}_t)]. \quad (3.29)$$

The importance, ι_n , of feature n is

$$\iota_n = \underbrace{v(s) - v(s^{-n})}_{\text{efficiency loss}} - \underbrace{(c(s) - c(s^{-n}))}_{\text{cost reduction}} + O(\|\Lambda\|^2). \quad (3.30)$$

We see that a feature is more important if it is an important contributor to the Markowitz portfolio (the first term in (3.30)) and if it is a persistent signal such that it reduces the turnover and, hence, the transaction costs (the second term in (3.30)).

This result provides intuition on economic feature importance based on an approximation. Empirically, we do not rely on this approximation but, instead, use ML tools to characterize the economic feature importance (3.27) as described in Section 5.3.

3 Benchmarks based on Standard Approaches

3.1 Standard Approach: Predicting Returns without T-Costs

The standard approach in the literature is to assume away transaction costs, that is, setting $\Lambda = 0$. In this case, the portfolio problem (3.9) becomes static in the sense that we can choose the optimal portfolio π_t at time t without regard to what happens at other time periods. Hence, the standard approach is focused on finding methods to predict returns and then using these return predictions to form a portfolio. For example, we can write the standard ML prediction problem as seeking to find a function f of stock characteristics $s_{n,t}$ that minimizes the mean-squared forecast errors for future 1-period (say, 1-month) excess returns $r_{n,t+1}$:

$$\min_{f: \mathbb{R}^K \rightarrow \mathbb{R}} \frac{1}{TN} \sum_{n,t} [r_{n,t+1} - f(s_{n,t})]^2. \quad (3.31)$$

This standard approach generates a function that approximates the conditional mean, $f(s_{i,t}) \cong E[r_{i,t+1}|s_{i,t}]$, when returns are stationary across time and assets, and the number of observations is large. The standard approach to turn such predictions into portfolio weights is to make factor, $\pi^{\text{factor-ML}}$, by going long a value-weighted average of the top 10% of the assets with the highest predicted returns $f(s_{i,t})$ while shorting the bottom 10% of the assets.

This simple factor approach ignores risk and transaction costs, but a more sophisticated method maximizes (3.9) while assuming zero transaction costs. Using the vector of expected excess returns $\mu(s_t) = (f(s_{1,t}), \dots, f(s_{N,t}))'$, the solution to (3.9) without transaction costs is

$$\pi_t^{\text{Markowitz-ML}} = \frac{1}{\gamma} \Sigma^{-1} \mu(s_t). \quad (3.32)$$

which is an ML-based version of the Markowitz portfolio.

3.2 Static Transaction Cost Optimization

A more sophisticated method is first to estimate the vector of expected excess returns $\mu(s_t)$ via (3.31) and then account for transaction costs in a second step. While this two-step procedure does not fully account for the dynamic nature of the problem, it serves as an interesting benchmark for our fully dynamic method. To see how this works, consider the problem of choosing an optimal portfolio π_t given the existing portfolio $g_t \pi_{t-1}$:

$$\max_{\pi_t \in \mathbb{R}^N} \left\{ \pi_t' \mu_t - \frac{\gamma}{2} \pi_t' \Sigma \pi_t - \frac{w}{2\phi} (\pi_t - g_t \pi_{t-1})' \Lambda (\pi_t - g_t \pi_{t-1}) \right\}. \quad (3.33)$$

Here, the transaction costs are divided by a “transaction-cost amortization parameter” ϕ to account for the static nature of the problem in an ad-hoc manner. A naive choice of this parameter is $\phi = 1$, which would mean that the objective (3.33) compares the returns earned over the next period (a “flow” variable) with the transaction costs (a “stock” variable) paid today. This comparison is problematic if the portfolio is expected to be held for many periods, so having the fudge factor ϕ is a simple way to address this problem. In particular, if the portfolio is expected to be held for $\phi = 6$ periods, we can amortize the trading cost over these six time periods, thus dividing the current transaction cost by six.

The solution to the static objective (3.33) is:

$$\begin{aligned} \pi_t^{\text{static-ML}} &= (\gamma \Sigma + \frac{w}{\phi} \Lambda)^{-1} (\mu_t + \frac{w}{\phi} \Lambda g_t \pi_{t-1}) \\ &= m^{\text{static}} g_t \pi_{t-1} + (I - m^{\text{static}}) \pi_t^{\text{Markowitz-ML}} \end{aligned} \quad (3.34)$$

where $m^{\text{static}} = (\gamma \Sigma + \frac{w}{\phi} \Lambda)^{-1} \frac{w}{\phi} \Lambda$. So we see that this strategy is a weighted average of the

inherited grown position, $g_t\pi_{t-1}$, and the current Markowitz portfolio. Said differently, this strategy always trades in the direction of the current Markowitz portfolio — so Markowitz is the “aim portfolio” in this static trading cost formulation. This solution is similar to our optimal portfolio with two exceptions. First, the aim portfolio in the fully dynamic model is more forward looking, distinguishing persistent signals from those with fast alpha decay. Second, the dynamic solution uses the utility optimal discount factor, m , rather than m^{static} .

4 Data and Empirical Methodology

4.1 Data and Inputs to the Portfolio Choice

Returns and Investment Universe

We use the dataset from Jensen et al. (2022b), a publicly available dataset and replication code of stock returns and characteristics, with the underlying return data sourced from CRSP and accounting data from Compustat.¹⁰ We restrict our sample to US common stocks (`shrcd`: 10, 11, and 12) traded on AMEX, NASDAQ, or NYSE (`exchcd`: 1, 2, or 3) with a market cap above the 50th percentile of NYSE stocks (denoted as large-cap stocks). For example, the group of large-cap stocks consists of the largest 1204 stocks at the end of 2020. This sample is deliberately conservative; the effects of trading cost optimization will be magnified among small, micro, and nano-cap stocks subject to notably larger trading costs. Our sample runs from 1952 to 2020, where the first part of the sample is used only for estimation, and our out-of-sample backtests run from 1981 to 2020.

Signals

To predict returns, covariances, and portfolio weights, we use 115 stocks characteristics (or features) studied in Jensen et al. (2022b).¹¹ We standardize each feature in each month by mapping the cross-sectional rank into the $[0,1]$ interval. We set missing values to 0.5 but require at least 57 non-missing features and non-missing market equity at the beginning of the month.

Investor Wealth and Optimization Methods

We assume that the investor wealth grows according to the realized market return, $w_t = w_{t-1}(1 + R_{m,t})$, such that the size of the investor is a stable share of the mar-

¹⁰The data, code, and documentation are available at <https://github.com/bkelly-lab/ReplicationCrisis/tree/master/GlobalFactors>.

¹¹Jensen et al. (2022b) studies 153 features. However, here we exclude features with poor coverage early in the sample. Table AI shows an overview of the features.

ket. This assumption means that the investor withdraws money when the portfolio has outperformed the market and vice versa when the portfolio has underperformed. For interpretability, we label each investor’s size by the corresponding wealth level by the end of 2020. In our baseline specification, the investor’s wealth evolves with the market return, so the final wealth by 2020 is \$10 billion.

We assume that the investor optimizes the portfolio each month using either Portfolio-ML, Multiperiod-ML, Static-ML, portfolio sort, or Markowitz-ML, as described below in Sections 4.2–4.3. These portfolio choice methods depend on trading costs and risks, which we estimate each month as described next. While we re-estimate trading costs and risks each month, the investor behaves as if trading costs and risks are constant over time. This assumption simplifies the ML problem, and while it may hurt out-of-sample performance that trading costs and risk do change over time, we find that the methods perform well nevertheless.

Trading Cost Matrix

Trading cost measured in dollars are given $TC_t = \frac{1}{2}\tau_t'\Lambda\tau_t$ for any vector of dollar trades, τ_t . In our empirical analysis, we let the trading cost matrix be diagonal and calibrate it based on the estimates in Frazzini et al. (2018). Specifically, we assume that the (expected and realized) market impact, $\frac{1}{2}\Lambda\tau_t$, is 0.1% when trading 1% of the daily dollar volume in a stock. This assumption means that the i^{th} diagonal entry in Λ_t , denoted $\Lambda_{i,t}$, satisfies $0.001 = \frac{1}{2}\Lambda_{i,t}0.01V_{i,t}$, which means that

$$\Lambda_{i,t} = \frac{0.2}{V_{i,t}}, \quad (3.35)$$

where $V_{i,t}$ is the expected daily dollar volume of stock i at time t . For example, trading \$5 million over a day in a stock with a daily volume of \$500 million moves the price by $\frac{1}{2}\frac{0.2}{\$500\text{m}} \times \$5\text{m} = 0.1\%$, leading to a transaction cost of $\frac{1}{2}\frac{0.2}{\$500\text{m}} \times (\$5\text{m})^2 = \$5000$. We follow Frazzini et al. (2018) and assume that the expected daily volume is equal to the average daily dollar volume over the last six months.

Variance-Covariance Matrix

We need to estimate the variance-covariance matrix, $\Sigma_t = \text{Var}_t(r_{t+1})$, at each time in a way that guarantees it to be positive definite and is broadly consistent with our estimates of expected returns. We use a factor model similar to the MSCI Barra risk model to accomplish these goals. Specifically, security characteristics are used as observable factor loadings, and latent factor returns are estimated via a simple regression (MSCI Barra,

2007).¹² Specifically, each trading day, we estimate a cross-sectional regression of stock returns on stock characteristics

$$r_{i,t+1} = S'_{i,t} \hat{f}_{t+1} + \epsilon_{i,t+1}, \quad (3.36)$$

and the regression coefficients, \hat{f}_{t+1} , are the estimated factor returns. Here, the observed characteristics, $S_{i,t}$, consist of a constant (the number one) and the 13 cluster characteristics from Jensen et al. (2022b). These cluster factors capture the main features of return predicting factors but do so in a simplified way to have a tractable variance-covariance matrix (simplified by reducing more than 100 factors to 13 clusters and by using a linear factor model rather than ML). Specifically, each stock's cluster characteristic, $S_{i,t}$, is its average rank of the characteristics in the cluster, standardized by subtracting the mean and dividing by the standard deviation each month. This standardization implies that the associated factors are long-short and dollar neutral, and, at the same time, the constant corresponds to an equal-weighted market factor. This structure means that the variance-covariance matrix is:

$$\hat{\Sigma}_t = S_t \text{Var}_t(\hat{f}_{t+1}) S'_t + \text{diag}(\text{Var}_t(\hat{\epsilon}_{i,t+1})). \quad (3.37)$$

Here, $\text{Var}_t(\hat{f}_{t+1})$ is estimated as the exponentially-weighted sample covariance matrix of factor returns over the past ten years of daily observations. We weight observations with exponential decays to put more weight on recent observations and, since correlations move slower than variances, we use a half-life of 378 days for correlations and 126 days for variances.¹³

Lastly, each stock's idiosyncratic variance, $\text{Var}_t(\hat{\epsilon}_{i,t+1})$, is estimated using an exponentially-weighted moving average of squared residuals, $\epsilon_{i,s}$, from (3.36) with a half-life of 126 days. We require at least 200 non-missing observations within the last 252 trading days. We use the median idiosyncratic variance within size groups to impute missing observations for stocks with less than 200 valid observations.

¹²The procedure of fixing factor loadings and estimating factor returns differs from models such as Fama and French (1993c), that fix factor returns and estimate loadings.

¹³Specifically, observations j days from t gets a weight of $w_{t-j} = c0.5^{j/\text{half-life}}$ where c is a constant ensuring that the sum of the weights is one.

4.2 Machine Learning Methodology

Machine Learning via Random Fourier Features

We use the machine learning method called random feature (RF) regression from Rahimi and Recht (2007).¹⁴ To understand the intuition behind this method, note that any function $f(s_t)$ can be approximated as

$$f(s_{i,t}) \approx RF(s_{i,t})\beta, \quad (3.38)$$

where $\beta \in \mathbb{R}^p$ is a vector of parameters and RF consists of random features. The RF method transforms the original features using random weights and a non-linear activation function. There are several ways to generate random features. We use so-called random Fourier features, which essentially approximate a function via its Fourier transformation.¹⁵ While this may sound complicated, it is straightforward to do in practice. We first simply draw some random Normal vectors, $w^j \in \mathbb{R}^{115} \sim iidN(0, \eta^2 I)$ for $j = 1, \dots, p/2$. Then, for each j , we create a pair of new features, $\sin(s'_{i,t}w^j)$ and $\cos(s'_{i,t}w^j)$, where the sine and cosine functions can capture non-linearities. We finally collect all these p random features:

$$RF(s_{i,t}) = [\sin(s'_{i,t}w^1), \cos(s'_{i,t}w^1), \dots, \sin(s'_{i,t}w^{p/2}), \cos(s'_{i,t}w^{p/2})]'$$

The RF method thus involves a vector of parameters β , estimated via a ridge regression, and two hyper-parameters, namely the number of random features p and the standard deviation of the random weights, η . We describe below how we choose these hyper-parameters via tuning.

Machine Learning about Expected Returns: Multiperiod-ML

We estimate expected returns using a ridge regression on the RF-transformed features. The resulting model can be viewed as a two-layer neural network with non-optimized weights in the first layer (the random features) and optimized weights in the final layer (the betas). We predict returns over three different horizons. The first model predicts excess returns over month $t + 1$, the second model predicts the average excess return over month $t + 2$ to $t + 6$, and the third model predicts the average excess return over month $t + 7$ to $t + 12$.

¹⁴See Kelly et al. (2022) for a detailed analysis of the theoretical properties of this RF methodology in the context of return prediction.

¹⁵The approach we use to generate random features is motivated by Sutherland and Schneider (2015), who find that it is preferable to alternative schemes with Gaussian weights.

Machine Learning Directly about the Optimal Portfolio: Portfolio-ML

Our Portfolio-ML learns about the aim portfolio via the relation

$$A_t = f(s_t)\beta = \text{diag}\left(\frac{1}{\sigma_{i,t}}\right)RF(s_t)\beta, \quad (3.39)$$

where $\beta \in \mathbb{R}^K$ is a parameter, we scale each asset's position by its volatility, $\sigma_{i,t} = \sqrt{\Sigma_{t,ii}}$, and RF consists of random Fourier features as described above. Note the objective for estimation is no longer return prediction but utility maximization, and the solution is given in Proposition 16.

4.3 Portfolio Tuning

The empirical implementation relies on several hyper-parameters as summarized in Table I. Consider first how we tune our Portfolio-ML method. This method runs a ridge regression on RF-transformed features, so we need to find the ridge parameter λ , the number of random features p , and the standard deviation of random weights η , collected in $h = (\lambda, p, \eta)$.

We tune h as follows. For each h , we compute a “validation backtest” in each year starting in 1971. Specifically, in each year $y \geq 1971$, we compute the optimal beta (3.26) for that h using monthly data from 1952 to $y - 1$. Using this beta, we compute the optimal portfolio for each month in year y and repeat this process each year until the end of our sample. This process creates – for each h – a backtest from 1971 onwards. These validation backtests are out-of-sample with respect to beta, but we still need to pick h .

Our “actual backtest” starts in 1981. Each year from 1981 onwards, we pick the hyper-parameter h with the highest realized utility in the validation backtest up until now (i.e., from 1971 until the previous year). Using this h and the corresponding beta, we compute the optimal portfolio over the next year, which is, therefore, truly out-of-sample with respect to both h and beta.

For the other methods (Multiperiod-ML, Static-ML, Markowitz-ML, portfolio sort), we first fit a model that predicts returns and then compute the optimal portfolio. We predict returns using a similar ML method based on random features using the tuning parameters h shown in Table I.

We show in the next section that Portfolio-ML outperforms Multiperiod-ML and Static-ML. In fact, the latter methods, in fact, deliver negative utility to the investor out of sample. This disappointing performance happens even though the ML model to predict returns works reasonably *well* in terms of how it ranks stocks. The problem is that the resulting portfolios tend to be poorly scaled because out-of-sample returns and risks for optimized portfolios do not match the scale of their ex-ante expected versions.

Table I. Hyper-Parameters

Hyper-parameter	Method		
First tuning layer, h	Portfolio-ML	Multiperiod-ML	Static-ML
Ridge penalty, λ	$\{0, e^4, e^5, \dots, e^8\}$	$\{0, e^{-10}, e^{-9.8}, \dots, e^{10}\}$	$\{0, e^{-10}, e^{-9.8}, \dots, e^{10}\}$
#random features, p	$\{2^6, 2^7, 2^8, 2^9\}$	$\{2^1, 2^2, \dots, 2^{10}\}$	$\{2^1, 2^2, \dots, 2^{10}\}$
Std of weights, η	$\{e^{-3}, e^{-2}\}$	$\{e^{-4}, e^{-3}, e^{-2}, e^{-1}\}$	$\{e^{-4}, e^{-3}, e^{-2}, e^{-1}\}$
Second tuning layer, h^*		Multiperiod-ML*	Static-ML*
Adjustment to mean, u		$\{0.25, 0.50, 1.00\}$	$\{0.25, 0.50, 1.00\}$
Adjustment to variance, v		$\{1, 2, 3\}$	$\{1, 2, 3\}$
Adjustment to t-cost, k		$\{1, 2, 3\}$	$\{\frac{1}{1}, \frac{1}{3}, \frac{1}{5}\}$

Note: The table shows the hyper-parameter space we use for portfolio tuning. For Portfolio-ML, λ is a ridge penalty, p is the number of random features, and η is the standard deviation of random weights. For Multiperiod-ML* and Static-ML* we add a second tuning layer; u shrinks the expected return vectors as $E_t[r_{t+\tau}]^* = uE_t[r_{t+\tau}]$, v increases stock variances as $\Sigma_t^* = \Sigma_t + v \text{diag}(\sigma_t)$, and k controls trading cost as $\Lambda_t^* = k\Lambda_t$.

So, this finding already shows the power of the Portfolio-ML method, namely its focus on the economic objective and on directly choosing portfolio weights, which immediately leads to an appropriate scaling of the portfolio with strong performance.

Nevertheless, we want to give the other methods a chance to compete with the Portfolio-ML method. To improve these alternative methods, we add an additional layer of portfolio tuning to Multiperiod-ML and Static-ML, where we add three additional tuning parameters: u, v, k . In particular, u shrinks the expected return vector towards zero, v increases the diagonal of the covariance matrix, and k increases the trading cost matrix:

$$\begin{aligned}
 E_t^*[r_{t+\tau}] &= uE_t[r_{t+\tau}], \\
 \Sigma_t^* &= \Sigma_t + v \text{diag}(\sigma_t), \\
 \Lambda_t^* &= k\Lambda_t.
 \end{aligned}
 \tag{3.40}$$

To estimate two layers of hyper-parameters, we proceed in the following way (which is rather involved, but, again, Portfolio-ML avoids this complexity). In the first layer, we produce a time series of out-of-sample expected returns based on $h = (\lambda, p, \eta)$ and, in the second layer, we produce optimal portfolios based on $h^* = (u, v, k)$. For the first layer, we update the RF models based on h each decade using the past 30 years of data. We estimate the random features models with each set of hyper-parameters over the first 20 years and pick the ones that lead to the lowest mean squared error over the last ten years. After finding the optimal hyper-parameters h , we re-train the model using all 30 years of data.

In the second layer of portfolio tuning, we update the hyperparameters h^* each year starting in 1981 by choosing the hyper-parameters that led to the highest utility since

1971. This two-layer approach is based on some experimentation to make these methods work, which gives these methods an advantage. Again, our main finding is that Portfolio-ML nevertheless performs even better. Figure A1 in the appendix shows the optimal parameters over time.

5 Empirical Results

This section reports our empirical results for each of our methods of portfolio choice. In our baseline specification, we consider an investor with a wealth of \$10 billion by the end of 2020 and a relative risk aversion of 10. We also consider other levels of wealth for comparison.

5.1 Out-of-Sample Portfolio Performance

Table II shows the out-of-sample performance for each method from 1981 to 2020. Judged by the performance before trading cost, the Markowitz-ML method is the clear winner with an impressive gross Sharpe ratio of 2.00. This finding shows that our methods for predicting risk and return perform well out-of-sample. The gross performance of the trading cost-aware portfolio choice methods (Portfolio-ML, Multiperiod-ML, Static-ML, Multiperiod-ML*, Static-ML*) is substantially lower than that of Markowitz-ML because these methods exploit fewer and less extreme trading opportunities to save on trading costs. Interestingly, several of these methods nevertheless realize a higher gross Sharpe ratio than the standard portfolio sort, presumably because they utilize information about risk and return.

After accounting for trading costs, the net return (and net Sharpe ratio) of the Markowitz-ML and portfolio sort methods are highly negative. Both methods trade too aggressively and are infeasible for the investor we consider. In contrast, the trading cost-aware methods still deliver positive net Sharpe ratios, reaching 1.38 for Portfolio-ML, an impressive performance given that we report out-of-sample results and account for the trading cost of a large investor with \$10 billion invested.

Turning to our main objective, which is to maximize the realized utility (3.12), Table II shows that the Portfolio-ML approach delivers the highest realized utility. In contrast, the methods in the top panel deliver negative realized utility. This top panel shows our results when all methods are fitted with a single layer of tuning (estimating portfolio weights or expected returns via random feature ML). While Portfolio-ML performs well with a single layer of tuning, the other methods deliver negative utility.

It is instructive to consider why Static-ML with a single layer of tuning delivers a negative utility despite its positive net Sharpe ratios. Figure 1.A shows that the indif-

Table II. Out-of-Sample Performance Statistics

Method	R	Vol.	SR _{gross}	TC	R-TC	SR _{net}	Utility	Turnover	Lev.
<u>One tuning layer</u>									
Portfolio-ML	0.20	0.14	1.43	0.008	0.19	1.38	0.095	0.32	3.60
Multiperiod-ML	0.32	0.34	0.95	0.182	0.14	0.41	-0.437	1.47	12.70
Static-ML	0.28	0.27	1.06	0.033	0.25	0.94	-0.106	0.76	11.21
Portfolio Sort	0.17	0.15	1.10	1.972	-1.81	-11.87	-1.921	2.60	2.00
Markowitz-ML	3.12	1.56	2.00	+	-	-	-	56.33	53.15
<u>Two tuning layers</u>									
Multiperiod-ML*	0.11	0.08	1.33	0.014	0.09	1.16	0.060	0.40	2.50
Static-ML*	0.13	0.10	1.36	0.024	0.11	1.11	0.060	0.61	3.22

Note: The table shows the out-of-sample performance of the various portfolio choice methods, rebalanced monthly from 1981–2020. Here, R is excess return; Vol. is volatility, SR_{gross} is the Sharpe ratio before trading cost; TC is trading cost, R-TC is excess return minus trading cost; SR_{net} is the Sharpe ratio after trading cost; Utility is the realized utility computed as the excess return after trading cost minus one-half times the assumed risk aversion of 10 times the realized portfolio variance; Turnover is the monthly average of the sum of absolute changes in portfolio weights, and Lev. is the portfolio leverage computed as the monthly average of the sum of absolute portfolio weights. All items except turnover and leverage are annualized. Portfolio-ML and Multiperiod-ML are the two dynamic trading cost optimization methods motivated by Proposition 13, Static-ML is the static trading cost optimization method from (3.34), Portfolio sort goes long/short the 10% of stocks with the highest/lowest 1-month expected return, and Markowitz-ML is the optimal portfolio absent trading cost from (3.32). For methods with one tuning layers, we search for the optimal hyper-parameters for a ridge regression implemented on RF-transformed features. For Multiperiod-ML* and Static-ML*, we add a second tuning layer to modify expected return, covariance, and trading cost inputs. An entry of “+” or “-” reflects an extremely high or low value.

ference curve corresponding to Static-ML goes below the origin, thus yielding a negative utility. This happens because this method realizes too high a risk relative to its ex-ante risk estimate. This is seen in Figure 1.A from the fact that the indifference curve crosses the frontier rather than being tangent (we note that the Static-ML frontier is not drawn but has a similar shape as that of Static-ML*). In other words, two factors determine the realized utility (i.e., the return net of trading costs and risk), out of sample: (i) how good the implementable efficient frontier is, and (ii) whether the method places the investor correctly on the frontier based on the investor’s risk aversion. While Static-ML produces a frontier that could deliver positive utility, it places the investor too far to the right on the frontier, thus realizing a negative utility.

To test Portfolio-ML even further, we also compare its performance to versions of the other methods where we give these other methods an extra “advantage” via a second layer of tuning as described in Section 4.3. This second layer of tuning is designed to improve the scaling of the portfolio, thus helping these methods to place the investor more correctly on the implementable efficient frontier.

Table II shows that the two-layer versions outperform the one-layer versions across all performance statistics. This result highlight that the second tuning layer is crucial for these methods based on ML about expected returns. Nevertheless, our Portfolio-ML

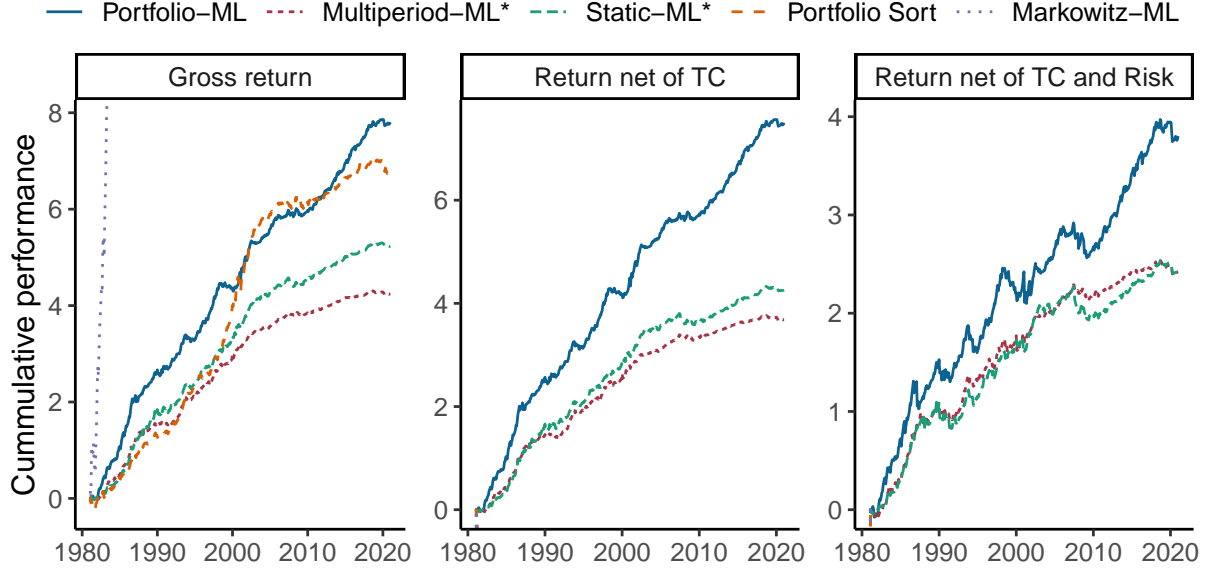


Figure 2. Performance over Time

Note: The left panel shows the cumulative sum of returns before trading cost, $r_{t+1}^{\pi, gross}$, for each portfolio method. The middle panel shows the cumulative sum of returns net of trading cost, $r_{t+1}^{\pi, net}$. The right panel shows the cumulative return net of trading cost (TC) and net of disutility from risk, computed as $r_{t+1}^{\pi, util} = r_{t+1}^{\pi, gross} - TC_t^\pi - \frac{\gamma}{2}(r_{t+1}^{\pi, net} - \bar{r}_{t+1}^{\pi, net})^2$, corresponding to the realized utility. We assume that the investors has a relative risk aversion of 10 and invested wealth of \$10 billion by the end of 2020.

continues to outperform these methods. This outperformance of Portfolio-ML relative to the two-layer methods shows a benefit of learning directly about portfolio weights, namely that the ML algorithm immediately searches for a well-scaled portfolio that delivers high utility – so no additional tuning layer is needed.

As seen from the notation in Table II, we add a superscript “*” to the implementations with two tuning layers. In the remainder of this section, we focus on the comparison between Portfolio-ML and the two-layers alternatives, Multiperiod-ML* and Static-ML*, studying their performance over time and the statistical significance of their performance differences.

Figure 2 shows that the outperformance of Portfolio-ML in terms of net returns and realized utility is consistent over time. The outperformance of Portfolio-ML is all the more remarkable when considering that the other methods were given the advantage of a second level of tuning to make them perform better. Figure 2 also shows some interesting time-series patterns in performance. For example, we see that several of these methods had a relatively lower performance during the dot-com bubble in 2000, the global financial crisis in 2008, and the COVID-19 crash in 2020.

One of the reasons behind the outperformance of Portfolio-ML is that this method keeps trading costs lower. This lower trading cost is achieved via a lower monthly turnover of 32% relative to 40% and 61% for Multiperiod-ML* and Static-ML*, respectively, as seen

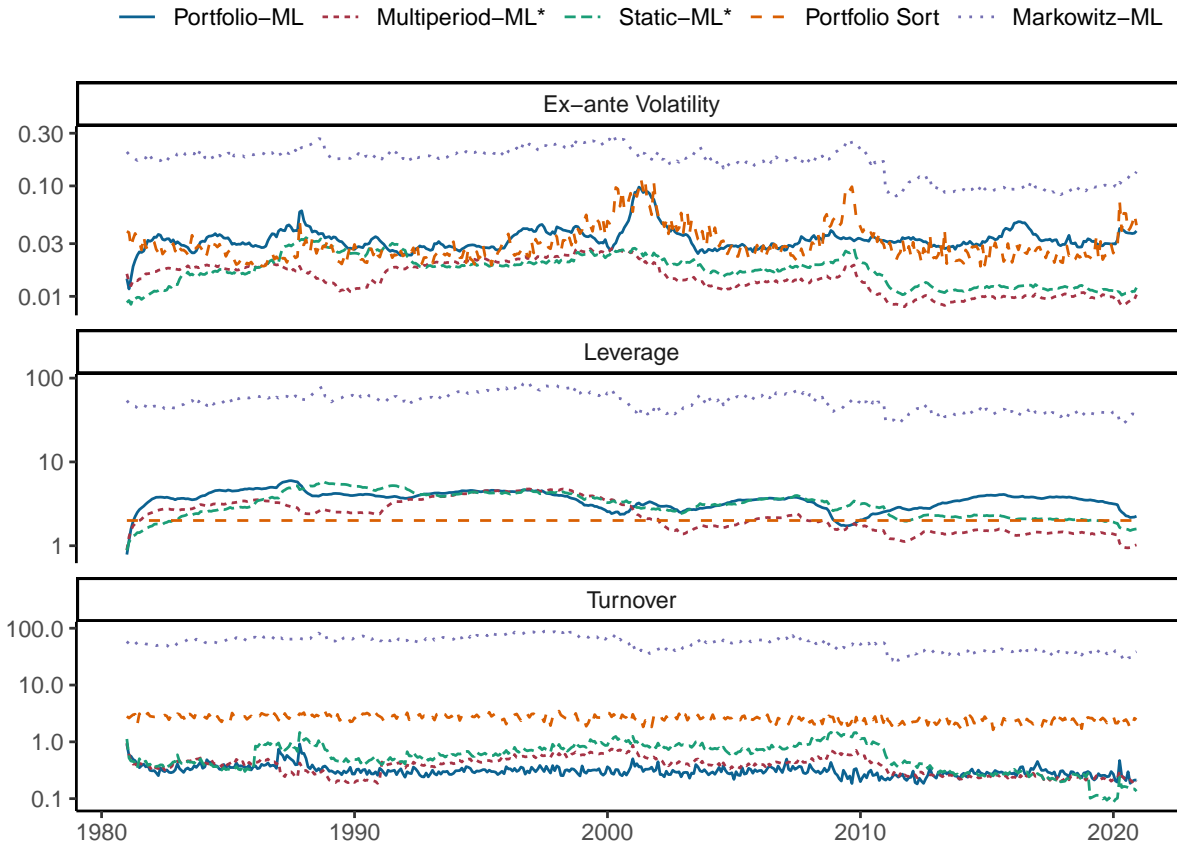


Figure 3. Portfolio Statistics over Time

Note: The top panel shows the ex-ante volatility of each portfolio method based on a monthly updated covariance matrix. The middle panel shows portfolio leverage defined as the sum of absolute portfolio weights. The lower panel shows the monthly portfolio turnover defined as $\sum_i |(\pi_t - g_t \pi_{t-1})_i|$, the sum of absolute differences between the current portfolio weight, π_t , and the grown portfolio weight from last month, $g_t \pi_{t-1}$. We use a logarithmic scale for the y-axis because of large differences across methods.

from Table II. Figure 3 shows how the ex-ante volatility, leverage, and turnover evolve over time, again showing that Portfolio-ML tends to have a lower turnover.

To visualize an example of some specific portfolio weights over time, Figure 4 depicts how the portfolio weights for Apply and Xerox stocks evolve for each method. Portfolio-ML adjusts its positions more slowly than the other methods, especially for the less liquid stock (Xerox).

Table III reports the statistical significance of the relative performance differences across portfolio choice methods. Specifically, the table reports the Bayesian probability that each method outperforms any of the other methods. To compute these pairwise probabilities of one method outperforming another, we first compute the utility flow (return net of trading costs and risk) of each method π at time $t + 1$ as $r_{t+1}^{\pi, util} = r_{t+1}^{\pi, gross} - TC_t^\pi - \frac{\gamma}{2}(r_{t+1}^{\pi, net} - \bar{r}_{t+1}^{\pi, net})^2$, where the relative risk aversion is $\gamma = 10$ as before. We then compute the utility difference between any two methods, say π and $\tilde{\pi}$, as $d_{\pi, \tilde{\pi}, t+1} =$

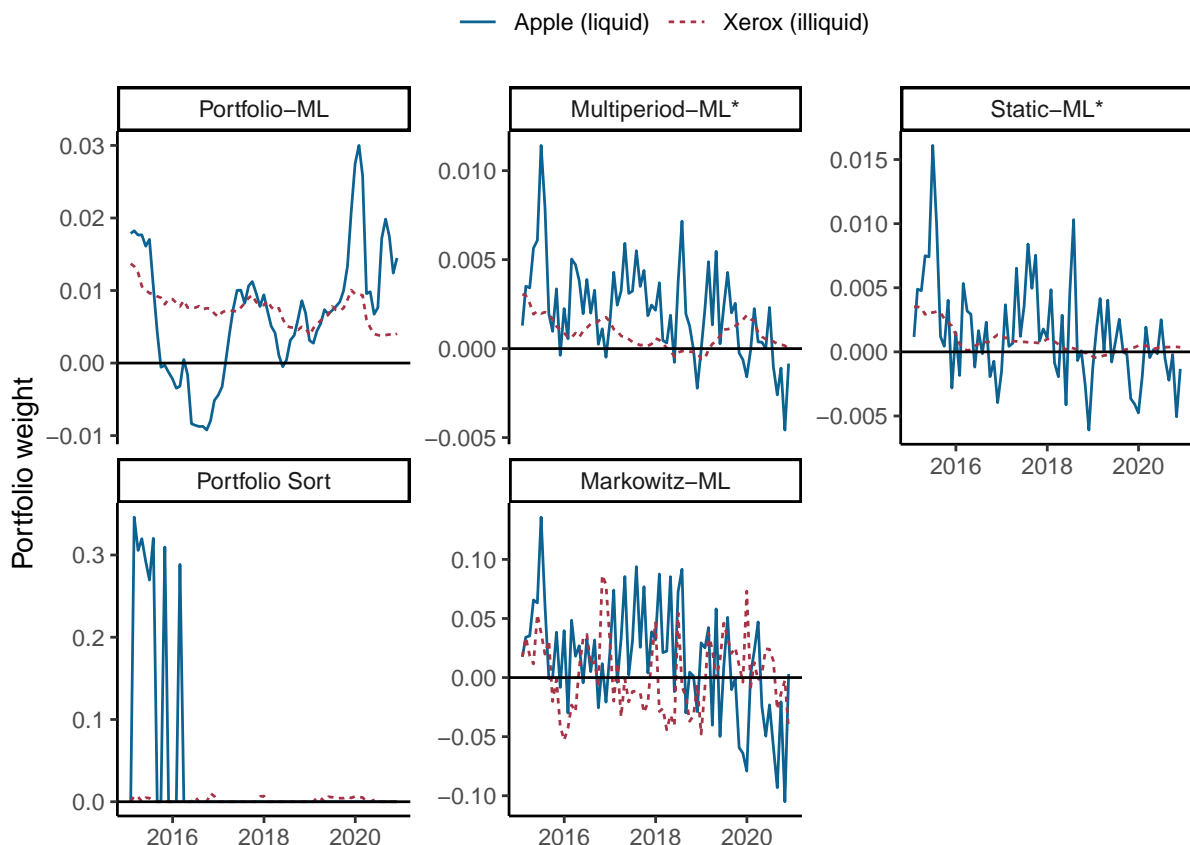


Figure 4. Portfolio Weights: Apple vs. Xerox

Note: The figure shows the portfolio weights of Apple and Xerox for each of the five portfolio choice methods, 2015–2020. Apple is chosen as an example of a relatively liquid stock and Xerox as a relatively illiquid stock over this time period. By the end of 2020, the average daily dollar volume over the past six months was \$16.39B for Apple and \$0.06B for Xerox.

$r_{t+1}^{\pi, util} - r_{t+1}^{\tilde{\pi}, util}$. The posterior of the true utility difference is then normally distributed with mean $\bar{d}_{\pi, \tilde{\pi}} = \frac{1}{T} \sum_{t=1}^T d_{\pi, \tilde{\pi}, t}$ and variance $\frac{1}{T-1} \sum_{t=1}^T (d_{\pi, \tilde{\pi}, t} - \bar{d}_{\pi, \tilde{\pi}})^2$ assuming that the difference is normally distributed with a non-informative prior about the mean and a known variance. Based on these calculations, Table III reports the posterior probability that $d_{\pi, \tilde{\pi}} > 0$, that is, the posterior probability that the first portfolio choice method, π , delivers a higher average utility than the second method, $\tilde{\pi}$.

Table III shows that the probability that Portfolio-ML delivers a higher expected utility than Multiperiod-ML*, Static-ML*, Portfolio sort, and Markowitz-ML are, respectively, 95%, 96%, 100%, and 100%, suggesting that the superiority of Portfolio-ML is not just random noise.

Alternatively, we can think of the probabilities in Table III as being approximately the p -value of a one-sided test that the realized utility of i is greater than j . Hence, we see that we can reject that Portfolio-ML delivers a lower realized utility than the other methods at conventional levels of significance.

Table III. Relative Probability of Outperformance

	Portfolio-ML	Multiperiod-ML*	Static-ML*	Portfolio Sort	Markowitz-ML
Portfolio-ML		95%	96%	100%	100%
Multiperiod-ML*	5%		51%	100%	100%
Static-ML*	4%	49%		100%	100%
Portfolio Sort	0%	0%	0%		100%
Markowitz-ML	0%	0%	0%	0%	

Note: The table shows the probability of the row method having a higher average utility than the column method. The probability is computed via an uninformative prior, assuming that the difference in utility is normally distributed. The utility flow of any method π at time $t + 1$ is $r_{t+1}^{\pi, util} = r_{t+1}^{\pi, gross} - TC_t^\pi - \frac{\gamma}{2}(r_{t+1}^{\pi, net} - \bar{r}_{t+1}^{\pi, net})^2$. One can also think of each number as the p -value in the test of whether the average utility of the portfolio choice method in the row is greater than the average utility of the method in the column.

Lastly, Table IV reports the return correlations of the various portfolio choice methods. We see that all methods are positively correlated, but the correlations tend to be modest in size. In addition to showing the relative connection across these methods of portfolio choice, these findings may also be informative about asset pricing more generally. Indeed, the Markowitz-ML portfolio return can be viewed as an estimate of the minimum-variance stochastic discount factor (SDF) in a frictionless market as shown by Hansen and Jagannathan (1991). Since risk adjustments depend on covariance with the SDF, a natural question is how closely SDF aligns with the corresponding measure designed for a market with frictions. The correlation between Portfolio-ML and Markowitz-ML is only 0.17, indicating that the marginal utility of an investor with \$10 billion using Portfolio-ML could be very different from risk adjustments in a frictionless market.

In summary, this section shows that Portfolio-ML outperforms the other methods, delivering a high net Sharpe ratio and high utility. While this strong performance suggests that Portfolio-ML works well, a few words of warning are in order. First, while the net performance is extremely good in our simulation, real-world investors seeking to achieve this performance must often pay fees to an asset manager (e.g., a hedge fund) running such strategies and face other real-world complications, potentially reducing performance. Second, investors might not have been able to realize this performance in real-time due to more limited computing power and a less developed ML methodology in the early sample.

Table IV. Portfolio Correlations

	Portfolio-ML	Multiperiod-ML*	Static-ML*	Portfolio Sort
Multiperiod-ML*	0.51			
Static-ML*	0.55	0.80		
Portfolio Sort	0.24	0.46	0.53	
Markowitz-ML	0.17	0.50	0.59	0.56

Note: The table shows the time-series correlation of the returns before trading costs for the various portfolio choice methods, 1981-2020.

In any case, this warning applies to any simulation, and the statistically significant out-performance of Portfolio-ML relative to other methods is an encouraging apples-to-apples test.

5.2 Evidence on the Implementable Efficient Frontier

Textbooks and real-world investors often depict their investment opportunities in terms of the achievable combinations of risk and expected return. This illustration highlights that investors seek a portfolio on the efficient frontier with the highest expected return for any level of risk. The textbook version of the efficient frontier – without trading costs – is a straight-line tangent to the hyperbola of risky investments. However, we propose that investors should focus on what we call the implementable efficient frontier, namely the efficient frontier net of trading costs.

Figure 1 illustrates our estimated implementable efficient frontier, out-of-sample. To understand how we have generated this plot, we start by describing the two benchmarks for a world without trading costs. The hyperbola is a mean-variance frontier of risky assets inspired by Markowitz (1952, 1959). We generate the points on the frontier by minimizing variance for a given mean and requiring that portfolio weights sum to 1:

$$\begin{aligned} \min_{\pi_t \in \mathbb{R}^N} \quad & \pi_t' \Sigma_t \pi_t, \\ \text{s.t.} \quad & \pi_t' \mu_t = k, \\ & \pi_t' \mathbf{1}_N = 1, \end{aligned}$$

where k is the required mean return, and $\mathbf{1}_N$ is a vector of ones. The solution is given by

$$\pi_t = \frac{c_t k - b_t}{d_t} \Sigma_t^{-1} \mu_t + \frac{(a_t - b_t k)}{d_t} \Sigma_t^{-1} \mathbf{1}_N, \quad (3.41)$$

where $a_t = \mu_t' \Sigma_t^{-1} \mu_t$, $b_t = \mathbf{1}'_N \Sigma_t^{-1} \mu_t$, $c_t = \mathbf{1}'_N \Sigma_t^{-1} \mathbf{1}_N$, and $d_t = a_t c_t - b_t^2$ are constants. Implementing this solution for a range of k 's generates the hyperbola. A standard presentation of the frontier uses one cross-section of stocks (i.e., one μ_t and one Σ_t) and presents the ex-ante expected frontier. In contrast, we show the realized frontier out-of-sample. Specifically, for each k , we update portfolio weights each month using (3.41). We then record the realized return and volatility before trading cost over the sample, 1981-2020. In contrast to the standard textbook presentation, the efficient frontier of risky assets in Figure 1 also accounts for out-of-sample performance decay. As such, it gives a more realistic picture of what investors could achieve absent trading costs.

The second benchmark for the case without trading cost is the Markowitz-ML portfolio. In a standard presentation, the line from this portfolio would be tangent to the

hyperbola. However, because our analysis is out-of-sample, this outcome is not ensured. In fact, the Markowitz-ML portfolio lies above the hyperbola.¹⁶

Next, we shift the attention from the frictionless benchmarks to our main focus, namely the implementable efficient frontier. These are computed via equation (3.11). Figure 1.A shows the efficient frontier after trading cost for an investor using different portfolio methods. We implement each method assuming an investor with a wealth that reaches \$10b by 2020.

We illustrate the performance of Markowitz-ML and portfolio sort with trading cost by scaling each portfolio to ex-post volatilities ranging from 0 to 10% in increments of 1%. We see their returns after trading costs are negative, except at very low volatilities. Hence, implementable efficient frontiers of these standard methods show that an investor maximizes utility by putting almost all wealth into the risk-free asset, thus choosing to hardly trade on these standard methods.

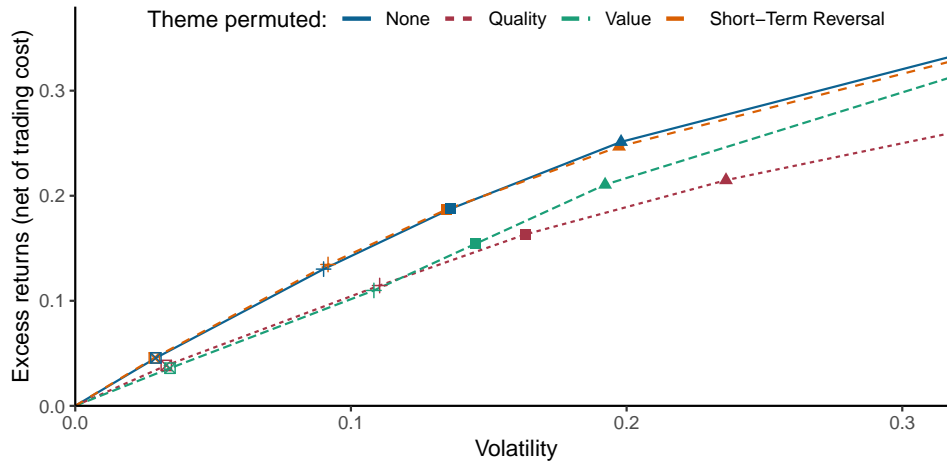
Turning to the trading-cost-aware methods, Portfolio-ML and Static-ML*, we see that their performance is much better. Instead of varying the ex-post volatilities, we implement the methods under five different relative risk aversions, $\gamma \in \{1, 5, 10, 20, 100\}$, and interpolate between their realized performance to plot an efficient frontier. Comparing the two implied frontiers, we see that Portfolio-ML leads to a higher achievable return for the same volatility. More generally, the figure shows that it is feasible for a large investor to generate an attractive implementable efficient frontier, even net of trading costs.

Panel B of Figure 1 shows how the implementable efficient frontier varies by investor size. Specifically, we implement the Portfolio-ML method for the same relative risk aversions as above, but now we also vary the investor wealth, $w_{2020} \in \{0, 10^9, 10^{10}, 10^{11}\}$. Such a plot is not interesting without trading costs since the efficient frontier is the same regardless of investor size. With trading costs, this is no longer the case. Price impact is increasing in trade size, so a larger investor must trade more slowly and focus more on liquid stocks. Naturally, these effects imply that larger investors have a worse risk-return tradeoff. The results in Figure 1.B quantifies how much worse. The figure shows the substantial cost of being a large investor. For example, an investor with \$10B and a relative risk aversion of 10 gets a net excess return of 19% at 14% volatility. If the

¹⁶This result might be surprising since the Markowitz-ML portfolio in a given period is proportional to the tangency portfolio of risky assets. Specifically, the tangency portfolio is $\pi^{\text{TPF}} = \frac{1}{b_t} \Sigma_t^{-1} \mu_t$, while the Markowitz-ML portfolio is $\pi_t^{\text{Markowitz-ML}} = \frac{1}{\gamma} \Sigma_t^{-1} \mu_t$. However, the scaling constant for Markowitz-ML, γ^{-1} , is fixed across time periods, while the scaling constant for the tangency portfolio, b_t^{-1} , varies significantly over time and even turns negative during a minority of periods. So the Markowitz portfolio can dominate the hyperbola for two reasons. First, when b_t is negative, the Markowitz portfolio shorts the tangency portfolio (which is on the lower branch of the hyperbola during such times), and the efficient frontier lies strictly above the hyperbola. Second, the differing time-varying scales of $\pi_t^{\text{Markowitz-ML}}$ and π^{TPF} (i.e., the different timing of the common underlying portfolio) turn out to work in favor of the Markowitz portfolio.

same investor had \$1B, a 14% volatility would provide a net return of 22%. In summary, once we introduce trading costs, we no longer have a unique, efficient frontier. Instead, the implementable efficient frontier depends on the investor’s size and portfolio choice method.

Panel A: Counterfactual Implementable Efficient Frontiers: With Trading Cost



Panel B: Counterfactual Efficient Frontiers: Without Trading Cost

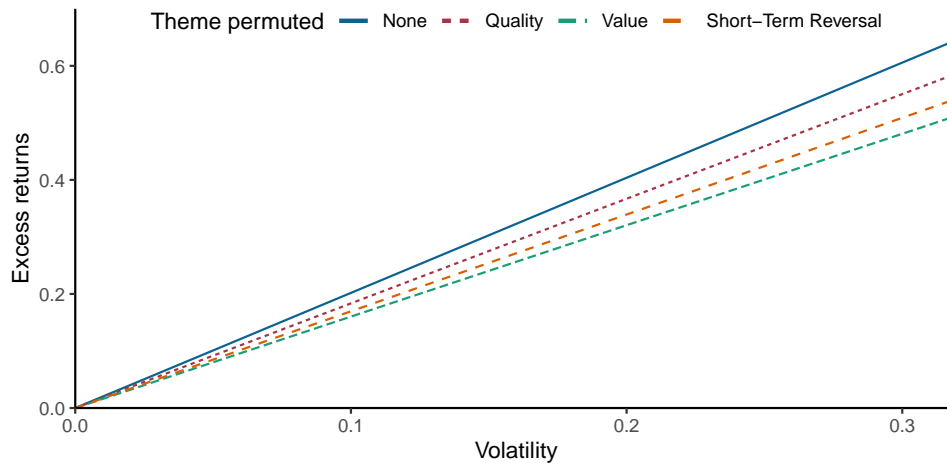


Figure 5. Feature Importance: Counterfactual Implementable Efficient Frontiers

Note: Panel A shows the implementable efficient frontier with trading costs for an investor with a wealth of \$10B by 2020. We implement the Portfolio-ML method using a counterfactual data set, where we permute all feature values related to either quality, value, or short-term reversal. The solid blue line shows the frontier using the actual data. Panel B shows the same analysis without trading cost, now using the Markowitz-ML method. In both panels, the relative risk aversions are 1 (circle), 5 (triangle), 10 (square), 20 (plus), and 100 (boxed cross) and the sample period is 1981-2020.

Finally, Figure 5 shows how the efficient frontier depends on access to certain features. We use a methodology known as permutation feature importance to assess this dependence. We provide a detailed description of this methodology in Section 5.3. Briefly, for a specific theme, say quality, we randomly shuffle (permute) all features related to this theme, effectively breaking their informational content. Using this counterfactual data,

we then implement Portfolio-ML (Panel A) or Markowitz-ML (Panel B). If the theme is important, breaking its informational content should lead to a worse risk-return tradeoff. In other words, destroying important features leads to a less desirable efficient frontier. We build separate counterfactual data sets by permuting quality, value, and short-term reversal signals. The solid blue line shows the efficient frontier with the original data.

Panel A shows the impact on the efficient frontier generated by Portfolio-ML, after trading cost, for an investor with \$10B in wealth. We see that quality and value signals are crucial for the implementable frontier. In contrast, destroying the informational content in short-term reversal signals barely changes the achievable frontier because these trading-cost-aware portfolio choice methods hardly use this signal anyway.

Panel B shows the impact on the efficient frontier generated by Markowitz-ML before trading cost. Here, all three themes are important. Interestingly, while short-term reversal has a minor effect on the implementable frontier with trading costs in Panel A, it greatly impacts the frontier without trading costs in Panel B.

In summary, looking at the efficient frontier without trading cost suggest that value, quality, and short-term reversal signals are all important for the efficient frontier. However, looking at the implementable efficient frontier of a large investor, value and quality remain important while short-term reversal does not. These results highlight that feature importance can change drastically when accounting for trading costs, and we explore this finding further in the next section.

5.3 Economic Feature Importance

Following the theoretical discussion in Section 2.3, we define economic feature importance as the drop in realized utility when excluding a feature from the information set of the investor. To implement this idea, we use the concept of permutation feature importance, introduced by Breiman (2001), which is a standard model-agnostic method for assessing feature importance of machine learning models (Molnar, 2022). The basic idea behind permutation feature importance is to permute features randomly and assess the decline in a user-defined value function. As such, we do not exclude the feature from the information set but destroy any predictive relationship between the feature and the outcome variable.

To compute economic feature importance for any portfolio choice method i , say Portfolio-ML, we first compute the baseline realized utility, $utility[\pi^i(s_{orig})]$, of the portfolio, π^i , computed based on the original features, s_{orig} . Next, for each feature j , we randomly permute its associated values at each given time while keeping all other features at their actual values. We then implement the portfolio method using the same parameters as in the original specification, but now with the permuted features, $s_{perm,j}$, as inputs. Finally, we compute the economic feature importance as the resulting drop in

utility, $FI_j^i = \text{utility}[\pi^i(s_{\text{orig}})] - \text{utility}[\pi^i(s_{\text{perm},j})]$. In other words, a feature j is economically important for method i if destroying its informational content leads to a large drop in realized utility, that is, a large FI_j^i .

A potential issue with permuting each feature separately is that substitution effects can distort the inference. For example, we include many different value features, so the effect of permuting a specific feature, such as book-to-market, is muted because the method can rely on other value features, such as assets-to-market or earnings-to-price. To handle substitution effects, we permute *all* features within a specific theme (or cluster) and record feature importance at the theme level. We use the 13 themes from Jensen et al. (2022b), shown in Table AI in the appendix.

Figure 6 shows the features' importance for three different methods. The left and middle panels show feature importance after trading costs for a large investor with a wealth of \$10b by 2020 for, respectively, Portfolio-ML and Multiperiod-ML*. The right panel shows feature importance without trading costs for Markowitz-ML, where we ignore trading costs because this method does not work after trading costs, making it meaningless to discuss net-of-cost feature importance. As such, the right panel serves as the benchmark of a frictionless market.

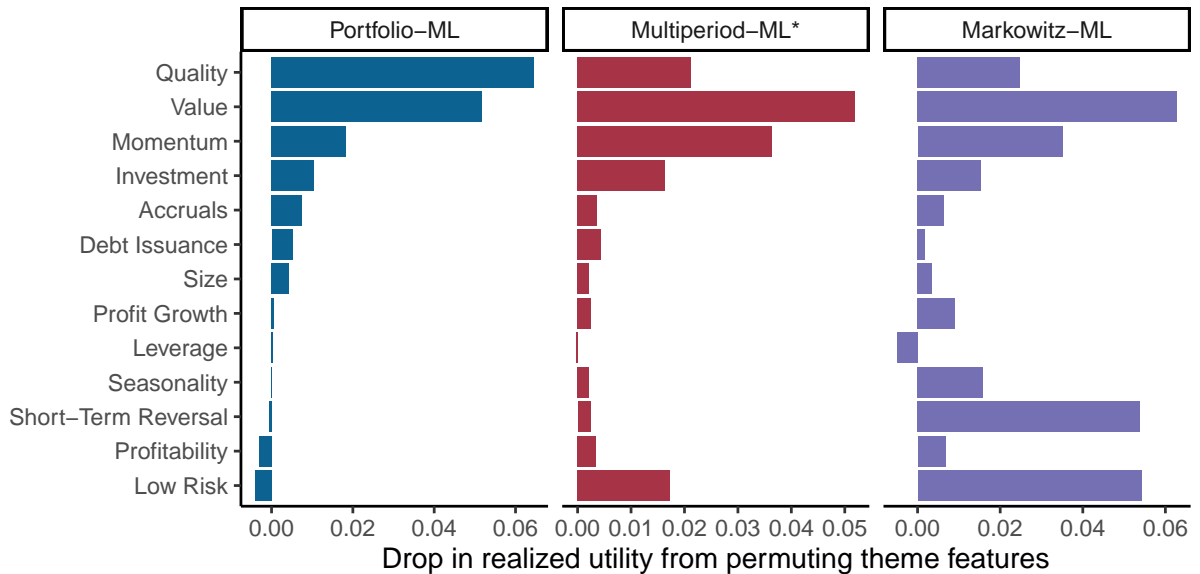


Figure 6. Economic Feature Importance

Note: The figure shows a utility-based feature importance measure for Portfolio-ML and Multiperiod-ML*, the two trading cost-aware methods motivated by Proposition 13, and Markowitz-ML, which is the optimal solution absent trading cost from (3.32). We randomly shuffle the associated features for each theme while keeping all other features at their actual value. We then implement each method based on this counterfactual data and measure feature importance as the difference in realized utility relative to the implementation that uses the actual data. For Markowitz-ML, we assume that the investor can trade without incurring the trading cost.

Looking at the frictionless benchmark in the right panel of Figure 6, we see that the

important feature themes before trading costs are value, short-term reversal, and low risk. Turning to the net-of-cost feature importance in the left and middle panels, we see that value remains important for a large investor. In contrast, short-term and low-risk are far less important due to the high turnover of many factors within these themes (see figure A2). This finding is consistent with the theoretical results of Section 2.3, namely that high-frequency features are less important in the presence of trading costs.

For example, the short-term reversal theme includes the short-term reversal factor, which has a monthly autocorrelation of -0.04 . This autocorrelation is not just low but actually negative, giving rise to large portfolio turnover.¹⁷ For a large investor, the predictive ability of short-term reversal is not enough to overcome the cost of trading it. Similarly, the low-risk theme includes features such as the past-month volatility or the maximum return over the past month, where the predictive ability is short-lived. (We note that some of the factors in these themes have low turnover, e.g., market beta, and may individually have a meaningful feature importance, but our results indicate that the overall themes are less important net of trading costs.) In comparison, book-to-market has a monthly autocorrelation of 0.94, indicating that it is a highly persistent feature, thus economizing on trading costs.

For Portfolio-ML, quality emerges as the most important theme. The median monthly autocorrelation of quality features is 0.93, so the result is again consistent with the theoretical findings. Perhaps surprisingly, momentum, generally considered a “fast” signal, is one of the most important themes for a large investor. However, several momentum features actually do exhibit meaningful persistence; for example, 12-month return momentum has a monthly autocorrelation of 0.87. Furthermore, momentum and value are negatively correlated, which leads to less trading because the two signals offset each other.

In summary, value, quality, and momentum are the economically important feature themes for a large investor facing trading costs. Before trading costs, high-frequency signals such as short-term reversal are the most important.

6 Conclusion

We develop a bridge between ML and portfolio choice with trading costs. To accomplish this bridge, we solve the optimal portfolio problem with transaction costs when returns are predictable via an *arbitrary* function of security characteristics and then show how the solution can be computed in a tractable way via machine learning directly about portfolio weights.

To evaluate the usefulness of our method – and, in fact, any method of portfolio

¹⁷The autocorrelation of a feature is computed as the average autocorrelation across all stocks with at least five years of monthly observations.

choice – we propose that investors should focus on the implementable efficient frontier, not the standard cost-agnostic efficient frontier. We show empirically that our method expands the implementable efficient frontier relative to other methods of portfolio choice. In other words, we find significant out-of-sample gains from our method even relative to sophisticated and more highly parameterized alternatives. We also consider several comparative statics, showing how the implementable efficient frontier contracts for larger investors facing higher market impact costs.

Finally, the method implies a novel view of which securities are important. Indeed, while standard methods that ignore transaction costs focus on transient features that work well on paper for small stocks, our method naturally selects persistent features of economic importance.

7 Appendix

The appendix is organized as follows. Appendix 7.1 presents implementation details, including how to compute the discount factor m (section 7.1) and details on Multiperiod-ML (7.1).

Appendix 7.2 contains proofs, including a key technical lemma for verifying optimality of policies (7.2), properties of m used in the proofs (7.2), proofs of Propositions 12 and 13 (7.2), proofs of Propositions 14 and 15 (7.2), the optimality of Portfolio ML (7.2), and economic feature importance (7.2).

Appendix 7.3 contains further empirical information, including an overview of the security characteristics used empirically (7.3), the estimated hyper-parameters over time (7.3), and the autocorrelation of the features and their importance for different return prediction horizons (7.3).

7.1 Implementation Details

Computing the Discount Factor m

Lemma 4 *Suppose that Λ and Σ are both diagonal. Let $\Lambda^{-1/2}\Sigma\Lambda^{-1/2} = \text{diag}(q_{i,i})$ is diagonal, then there exists a unique diagonal solution \tilde{m} to (3.111) such that $\Lambda^{-1/2}\tilde{m}\Lambda^{1/2}\bar{g}$ has all eigenvalues below one in absolute value. It is given by*

$$m_{i,i} = \frac{2}{w^{-1}\gamma q_{i,i} + G_{i,i} + 1 + \sqrt{(w^{-1}\gamma q_{i,i} + G_{i,i} + 1)^2 - 4G_{i,i}}} \quad (3.42)$$

Proof of [. Proof of Lemma 4] The proof follows by direct calculation. ■

Another special case with a closed-form solution is when G has a rank of one, which is not a realistic case, but turns out to be a useful approximation. Specifically, we have that $G = \bar{g}\bar{g}' + \frac{1}{(1+r^f+\bar{\mu})^2}\Sigma \cong \bar{g}\bar{g}'$, where¹⁸ $\bar{g} = (1+r^f+\bar{\mu})^{-1}(1+r^f+E[\mu])$ and the approximation is based on the idea that \bar{g} is a vector of numbers close to one, whereas Σ is much smaller with numbers of the order of 0.10² when monthly volatility is around 10%.

Lemma 5 *Suppose that Λ is diagonal. In the case when $G = \xi\xi'$ for some vector $\xi > 0$, then the unique solution m to (3.99) such that $m \text{diag}(\xi)$ has all eigenvalues below one in absolute value and $\tilde{m} \in \mathcal{S}(0,1)$ given by $m = \Lambda^{-1/2}\tilde{m}\Lambda^{1/2}$ where*

$$\tilde{m} = \text{diag}(\xi)^{-1/2} 0.5(\hat{\Sigma} - (\hat{\Sigma}^2 - 4I)^{1/2}) \text{diag}(\xi)^{-1/2} \quad (3.43)$$

¹⁸We abuse the notation and use \bar{g} to denote both the vector and the diagonal matrix.

and $\hat{\Sigma} = \text{diag}(\xi)^{-1/2}(w^{-1}\Lambda^{-1/2}\gamma\Sigma\Lambda^{-1/2} + \text{diag}((\xi_i^2 + 1))) \text{diag}(\xi)^{-1/2}$ and $(\hat{\Sigma}^2 - 4I)^{1/2}$ is the unique positive-definite square root.

Proof of [Proof of Lemma 5] We have that (3.111) takes the form

$$\tilde{m} = \left(\Lambda^{-1/2}\Sigma\Lambda^{-1/2} + I + \Lambda^{-1/2} \text{diag}(\xi)(\Lambda^{1/2}(I - \tilde{m})\Lambda^{1/2}) \text{diag}(\xi)\Lambda^{-1/2} \right)^{-1}. \quad (3.44)$$

and the assumption of a diagonal Λ implies

$$\tilde{m} = \left(\Lambda^{-1/2}\Sigma\Lambda^{-1/2} + I + \text{diag}(\xi)(I - \tilde{m}) \text{diag}(\xi) \right)^{-1}. \quad (3.45)$$

We abuse the notation and use ξ to denote $\text{diag}(\xi)$. Let $\tilde{\Sigma} = \Lambda^{-1/2}\Sigma\Lambda^{-1/2} + I + \xi^2$. Then, (3.45) takes the form

$$\tilde{m} = \left(\tilde{\Sigma} - \xi\tilde{m}\xi \right)^{-1}. \quad (3.46)$$

Define

$$\hat{\Sigma} = \xi^{-1/2}\tilde{\Sigma}\xi^{-1/2} = \xi^{-1/2}\Lambda^{-1/2}\Sigma\Lambda^{-1/2}\xi^{-1/2} + \xi^{-1} + \xi > 2I,$$

where the last inequality follows because $\xi + \xi^{-1} \geq 2$ for any positive number ξ . Let also $\hat{m} = \xi^{1/2}\tilde{m}\xi^{1/2}$. Then, we get

$$\hat{m}^2 - \hat{\Sigma}\hat{m} + I = 0 \quad (3.47)$$

which has 2^N solutions. The smallest solution is given by

$$\hat{m} = 0.5(\hat{\Sigma} - (\hat{\Sigma}^2 - 4I)^{1/2}). \quad (3.48)$$

The function $f(x) = 0.5(x - (x^2 - 4)^{1/2}) = 2/(x + (x^2 - 4)^{1/2}) < 1$ for all $x > 2$, and the claim follows. \blacksquare

Starting with this approximation, we can compute the exact m stepwise, as follows. Note first that the set \mathcal{S} of symmetric, positive definite matrices is a partially-ordered set with respect to the positive semi-definite order: We say that $m_1 \leq m_2$ if $m_2 - m_1$ is positive semi-definite. Further, we let $\mathcal{S}(0, 1)$ be the set of positive semi-definite matrices with eigenvalues between zero and one.

Suppose for simplicity that Λ is diagonal. Since the optimum is unique, the proof of Proposition 12 implies that (3.111) has a unique solution $\tilde{m} \in \mathcal{S}(0, 1)$. Remarkably, this solution automatically satisfies the transversality condition $\bar{g}^{1/2}\tilde{m}\bar{g}^{1/2} \in \mathcal{S}(0, 1)$. The following lemma shows how to construct this unique solution.

Lemma 6 *The unique solution $\tilde{m} \in \mathcal{S}(0, 1)$ to (3.111) can be computed as follows: $m =$*

$\Lambda^{-1/2}\tilde{m}\Lambda^{1/2}$, where \tilde{m} can be found by iterating the mapping F :

$$F(\tilde{m}) = \left(w^{-1}\Lambda^{-1/2}\gamma\Sigma\Lambda^{-1/2} + I + \Lambda^{-1/2}((\Lambda^{1/2}(I - \tilde{m})\Lambda^{1/2}) \circ G)\Lambda^{-1/2} \right)^{-1}. \quad (3.49)$$

Indeed, F maps $\mathcal{S}(0, 1)$ into itself, is monotonic with respect to the positive semi-definite order. Furthermore, it has a unique fixed point $\tilde{m}_* \in \mathcal{S}(0, 1)$, and its iterations converge to this unique fixed point from any starting point m_0 in $\mathcal{S}(0, 1)$ satisfying either $m_0 \leq F(m_0)$ or $m_0 \geq F(m_0)$. In particular, it converges upward from the smallest starting point $0 \in \mathcal{S}(0, 1)$:

$$F(0) \leq F(F(0)) \leq \dots \leq F(\dots(F(0))) \rightarrow \tilde{m}_*$$

Furthermore, the map F is monotone decreasing in the matrix G in the sense of positive semi-definite order: $\tilde{m}(G_1) \leq \tilde{m}(G_2)$ whenever $G_1 \geq G_2$. In particular, if $G = \bar{g}\bar{g}' + \frac{1}{(1+r^f+\bar{\mu})^2}\Sigma$ with $\bar{g} = (1+r^f+\bar{\mu})^{-1}(1+r^f+E[\mu])$, and if Λ is diagonal, then it also converges downward from the starting point $\tilde{m}(\bar{g})$ from Lemma 5:

$$F(\tilde{m}(\bar{g})) \geq F(F(\tilde{m}(\bar{g}))) \geq \dots \geq F(\dots(F(\tilde{m}(\bar{g})))) \rightarrow \tilde{m}_* \quad (3.50)$$

so the first iterations of these provide lower and upper bounds:

$$\begin{aligned} & \left(w^{-1}\Lambda^{-1/2}\gamma\Sigma\Lambda^{-1/2} + \text{diag}((G_{i,i} + 1)) \right)^{-1} \\ & \leq \tilde{m}_* \leq \left(w^{-1}\Lambda^{-1/2}\gamma\Sigma\Lambda^{-1/2} + I + ((I - \tilde{m}(\bar{g})) \circ G) \right)^{-1}. \end{aligned} \quad (3.51)$$

Proof of T. The only claim that requires proof is the fact that $\tilde{m}(\bar{g}) \geq F(\tilde{m}(\bar{g}); G)$. Indeed, by the monotonicity of F in G we have

$$\tilde{m}(\bar{g}) = F(\tilde{m}(\bar{g}); \bar{g}\bar{g}') \geq F(\tilde{m}(\bar{g}); G),$$

and the claim follows. This inequality, combined with monotonicity, implies the required sequence of inequalities (3.50). ■

Implementation Details for Multiperiod-ML

Suppose that Λ is diagonal. While formula (3.21) requires an infinite sum, in our numerical implementation, we use the approximation

$$\begin{aligned}
A_t &= (I - m)^{-1}(I - (m \operatorname{diag}(\bar{g})))^2(I - (m \operatorname{diag}(\bar{g})))^{-2} \sum_{\tau=0}^{\infty} (m \operatorname{diag}(\bar{g}))^{\tau} c \Sigma^{-1} E_t[r_{t+1+\tau}] \\
&= (I - m)^{-1}(I - (m \operatorname{diag}(\bar{g})))^2 \sum_{\tau=0}^{\infty} (m \operatorname{diag}(\bar{g}))^{\tau} (I - (m \operatorname{diag}(\bar{g})))^{-2} c \Sigma^{-1} E_t[r_{t+1+\tau}] \\
&= (I - m)^{-1}(I - (m \operatorname{diag}(\bar{g}))) (I - (m \operatorname{diag}(\bar{g}))) \sum_{\tau=0}^{\infty} (m \operatorname{diag}(\bar{g}))^{\tau} \tilde{c} \Sigma^{-1} E_t[r_{t+1+\tau}] \\
&\approx (I - m)^{-1}(I - (m \operatorname{diag}(\bar{g})))^2 (I - (m \operatorname{diag}(\bar{g})))^{k+1}{}^{-1} \sum_{\tau=0}^k (m \operatorname{diag}(\bar{g}))^{\tau} \tilde{c} \Sigma^{-1} E_t[r_{t+1+\tau}]
\end{aligned} \tag{3.52}$$

where we have defined

$$\tilde{c} = (I - (m \operatorname{diag}(\bar{g})))^{-2} c. \tag{3.53}$$

7.2 Proofs

Properties of the Implementable Efficient Frontier

We note that the textbook efficient frontier is usually defined in terms of a risk-minimization problem rather than the return-maximization in (3.10). Hence, we could consider the corresponding definition of the implementable efficient frontier as the combination of volatilities and expected net returns, $(\sigma(k), k)_{k \geq 0}$, such that risk is minimal for that level of net return:

$$\sigma(k)^2 = \min_{\pi_t} E[\pi_t' \Sigma \pi_t] \quad \text{s.t.} \quad E[r_{t+1}^{\pi, net}] = k \tag{3.54}$$

However, this definition is less helpful for two reasons. First, no solution exists for large enough k . Second, this definition cannot produce the downward-sloping portion of the frontier seen in Figure 1.

Proof of [. Proof of Proposition 11] (i) We first show that the net Sharpe ratio is decreasing along the frontier. To see that, consider two risk levels, $\sigma_1 < \sigma_2$. Let π^2 be the frontier portfolio corresponding to σ_2 . This portfolio has the highest expected net return for this risk level and the highest net Sharpe ratio. If we scale down this portfolio to $\pi^1 = \frac{\sigma_1}{\sigma_2} \pi^2$ (putting the rest of the money in the risk-free asset), then the risk becomes $\frac{\sigma_1}{\sigma_2} \sigma^2 = \sigma_1$. Then we have:

$$\max_{\pi_t \in \Pi \text{ s.t. } E[\pi_t' \Sigma \pi_t] = \sigma_1^2} E[r_{t+1}^{\pi, net}] \geq E[r_{t+1}^{\pi^1, net}] > \frac{\sigma_1}{\sigma_2} E[r_{t+1}^{\pi^2, net}] \tag{3.55}$$

Here, the first inequality follows from the definition of the frontier as the maximum. The second inequality follows from the fact that gross returns are linear, but transaction costs are quadratic and $\frac{\sigma_1}{\sigma_2} < 1$. Dividing both sides of (3.60) by σ_1 , we see that the net Sharpe ratio is decreasing in σ .

(vi) Consider the implementable efficient frontier corresponding to a wealth of w_1 and w_2 , where $w_1 < w_2$. Take a point on the frontier of w_2 corresponding to the risk of σ and a portfolio π^2 . Then with wealth w_1 , the portfolio π^2 delivers a higher net return with the same risk (and there may exist another portfolio with an even higher net return for this level of wealth), so the frontier of w_1 must be above that of w_2 .

(ii)–(iii) To prove the concavity of the implementable efficient frontier, consider two risk levels, σ_1, σ_2 . Let $\pi(\sigma)$ be the frontier portfolio corresponding to σ , and let

$$R(\sigma) = \max_{\pi_t \in \Pi \text{ s.t. } E[\pi_t' \Sigma \pi_t] = \sigma} E[r_{t+1}^{\pi, net}] \quad (3.56)$$

Define

$$\tilde{\pi} = 0.5(\pi(\sigma_1) + \pi(\sigma_2)). \quad (3.57)$$

Then,

$$\begin{aligned} E[\tilde{\pi}' \Sigma \tilde{\pi}] &= E[0.25((\pi(\sigma_1)' \Sigma \pi(\sigma_1) + (\pi(\sigma_2)' \Sigma \pi(\sigma_2) + 2(\pi(\sigma_1)' \Sigma \pi(\sigma_2)))] \\ &\leq 0.25(\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2) = (0.5(\sigma_1 + \sigma_2))^2, \end{aligned} \quad (3.58)$$

where we have used a modified Cauchy-Schwarz inequality

$$\begin{aligned} E[(\pi(\sigma_1)' \Sigma \pi(\sigma_2))] &\leq E[((\pi(\sigma_1)' \Sigma \pi(\sigma_1))^{1/2} ((\pi(\sigma_2)' \Sigma \pi(\sigma_2))^{1/2})] \\ &\leq E[((\pi(\sigma_1)' \Sigma \pi(\sigma_1))^{1/2} E[(\pi(\sigma_2)' \Sigma \pi(\sigma_2))^{1/2}]] \leq \sigma_1 \sigma_2 \end{aligned} \quad (3.59)$$

Therefore,

$$\begin{aligned} 0.5(R(\sigma_1) + R(\sigma_2)) &= 0.5(E[r_{t+1}^{\pi(\sigma_1), net}] + E[r_{t+1}^{\pi(\sigma_2), net}]) \\ &= E[r_{t+1}' \tilde{\pi} - 0.5(TC^{\pi(\sigma_1)} + TC^{\pi(\sigma_2)})]. \end{aligned} \quad (3.60)$$

Since transaction costs are convex in π_t , we have

$$-0.5(TC^{\pi(\sigma_1)} + TC^{\pi(\sigma_2)}) \leq -TC^{\tilde{\pi}}.$$

Thus,

$$E[r_{t+1}^{\tilde{\pi}, net}] \geq 0.5(R(\sigma_1) + R(\sigma_2)),$$

while

$$\sigma(\tilde{\pi}) = (E[\tilde{\pi}' \Sigma \tilde{\pi}])^{1/2} \leq 0.5(\sigma_1 + \sigma_2).$$

Thus, we get $R(\sigma(\tilde{\pi})) \geq 0.5(R(\sigma_1) + R(\sigma_2))$, and hence, the required concavity follows if $R(\sigma)$ is increasing on $[0.5(\sigma_1 + \sigma_2), \sigma(\tilde{\pi})]$.

Now, pick a $\gamma > 0$. Then, clearly, π^γ belongs to the efficient frontier, corresponding to some $\sigma(\gamma)$: Otherwise, we could increase net expected return keeping the variance fixed. We also note that for large γ , the effect of transaction costs is negligible and, hence, the efficient frontier for $\sigma \approx 0$ approximately coincides with the frictionless one, and hence $k(\sigma)$ is monotone increasing for $\sigma \approx 0$.

The set of eligible portfolios (adapted, square-integrable processes) is a Hilbert space H we can define operators A, B and a vector $x \in H$ so that $E[r^{\pi, net}] = \langle x, \pi \rangle - 0.5\langle Ax, x \rangle$ and $E[\pi' \Sigma \pi] = \langle \pi B, \pi \rangle$, where $\langle \cdot, \cdot \rangle$ is the inner product in the Hilbert space. We consider finite-dimensional approximations of the quadratic problem and thus assume that A, B are finite-dimensional matrices. Then, the first order condition is

$$x - A\pi = \lambda B\pi \quad (3.61)$$

where λ is the Lagrange multiplier of the constraint $\langle \pi, B\pi \rangle = \sigma^2$. Now,

$$\pi = (A + \lambda B)^{-1}x, \quad (3.62)$$

and we need to solve the equation

$$\langle (A + \lambda B)^{-1}B(A + \lambda B)^{-1}x, x \rangle = \sigma^2 \quad (3.63)$$

For the increasing part of the frontier, the constraint $\langle \pi, B\pi \rangle \leq \sigma^2$ is binding and $\lambda > 0$. This defines $\sigma_*^2 = \langle A^{-1}BA^{-1}x, x \rangle$.

Beyond that we need to use the eigen-decomposition of $B^{-1/2}AB^{-1/2}$ and define $\tilde{x} = B^{-1/2}x$. Then, if ν_i are the eigenvalues of $BB^{-1/2}AB^{-1/2}$ and \tilde{x}_i are the coordinates of \tilde{x} in the eigen-basis, we get that we need to maximize

$$\begin{aligned} & \langle x, (A + \lambda B)^{-1}x \rangle - 0.5\langle (A + \lambda B)^{-1}(A + \lambda B - \lambda B)x, (A + \lambda B)^{-1}x \rangle \\ &= 0.5\langle x, (A + \lambda B)^{-1}x \rangle + 0.5\lambda\sigma^2 \\ &= 0.5 \sum_i \tilde{x}_i^2 (1 + \lambda\nu_i)^{-1} + 0.5\lambda\sigma^2 \end{aligned} \quad (3.64)$$

under the constraint

$$\sum_i \tilde{x}_i^2 (1 + \lambda\nu_i)^{-2} = \sigma^2. \quad (3.65)$$

This function (as a function of λ) explodes for $\lambda = -1/\nu_i$.

Now, for the mean-variance optimization problem, the solution is

$$\pi = (A + \gamma B)^{-1}x \quad (3.66)$$

and the variance

$$\langle (A + \gamma B)^{-1}x, B(A + \gamma B)^{-1}x \rangle \quad (3.67)$$

is monotone decreasing in γ and converges to σ_*^2 when $\gamma \rightarrow 0$. At the same time,

$$R(\gamma) = \langle (A + \gamma B)^{-1}x, x \rangle \quad (3.68)$$

is also monotone decreasing in γ . ■

Verification Lemma

Our proofs are based on the following auxiliary result.

Lemma 7 *For simplicity, we normalize $\gamma/w = 1$. Let $\bar{\Lambda}_t = E_t[g_{t+1}\Lambda g_{t+1}]$. For any solution m_t to*

$$m_t = (\Sigma + \Lambda + \bar{\Lambda}_t)^{-1} \left(E_t[g_{t+1}\Lambda m_{t+1}g_{t+1}]m_t + \Lambda \right), \quad (3.69)$$

define

$$N_{t,t+\tau} = \prod_{\tau=1}^{\theta} m_{t-\tau+1} g_{t-\tau+1} \quad (3.70)$$

and

$$\tilde{N}_{t,t+\tau} = \prod_{\tau=1}^{\theta} m_{t-\tau+1} \Lambda^{-1} g_{t-\tau+1} \Lambda. \quad (3.71)$$

Suppose that

$$\sum_{\tau=1}^{\infty} \|E_t[N'_{t,t+\tau} N_{t,t+\tau}]\|^{1/2} < \infty \quad (3.72)$$

and

$$\sum_{\tau=1}^{\infty} \|E_t[\tilde{N}'_{t,t+\tau} \tilde{N}_{t,t+\tau}]\|^{1/2} < \infty. \quad (3.73)$$

Define

$$c_t = m_t \Lambda^{-1} \Sigma \quad (3.74)$$

and

$$Q_t = E_t \left[\sum_{\tau=0}^{\infty} \tilde{N}_{t,t+\tau} c_{t+\tau} \text{Markowitz}_{t+\tau} \right] \quad (3.75)$$

and let

$$\pi_t = \pi(s_t, s_{t-}) = \sum_{\theta=0}^{\infty} N_{t-\theta, t} Q(s_{t-\theta}). \quad (3.76)$$

Then, all series converge in L_2 and π_t is optimal among all bounded stationary processes π_t . Furthermore, it satisfies the recursive relationship

$$\pi(s_t, s_{t-}) = Q(s_t) + m_t g_t \pi(s_{t-1}, s_{t-1-}) \quad (3.77)$$

Proof of [. Proof of Lemma 7] Due to the strict convexity of the objective, it suffices to verify the first order conditions. Let

$$\mathcal{O}(\pi) = \min_{\pi \in L_2} E \left[-2\mu(s_t)' \pi_t + \pi_t' \Sigma \pi_t + (\pi_t - g_t \pi_{t-1})' \Lambda (\pi_t - g_t \pi_{t-1}) \right] \quad (3.78)$$

Standard convexity arguments imply that it suffices to derive and verify the first order conditions for our candidate solution.

Let s_{t-} denote the history of s_t and $\pi_t = \pi(s_t, s_{t-})$ be a candidate optimal policy and Then, by direct calculation, using the ergodicity property, we get while the law of iterated expectations implies that

$$E[\pi_{t-1}' g_t \Lambda g_t \pi_{t-1}] = E[\pi_t' g_{t+1} \Lambda g_{t+1} \pi_t] = E[\pi_t' \hat{\Lambda}_t \pi_t],$$

where

$$\bar{\Lambda}_t = E_t[g_{t+1} \Lambda g_{t+1}]. \quad (3.79)$$

Therefore,

$$\begin{aligned} \mathcal{O}(\pi) &= E \left[-2\mu(s_t)' \pi(s_t, s_{t-}) + \pi(s_t, s_{t-})' (\Sigma + \Lambda + \bar{\Lambda}_t) \pi(s_t, s_{t-}) \right] \\ &\quad - 2E[\pi(s_t, s_{t-})' g_{t+1} \Lambda \pi(s_{t+1}, s_{t+1-})], \end{aligned} \quad (3.80)$$

In order to compute the first order conditions, we need to calculate the Frechet derivative of (3.80) with respect to π . To this end, we consider a small perturbation $\pi \rightarrow \pi + \varepsilon Y$ and calculate the first order term in ε , so that

$$\mathcal{O}(\pi_t + \varepsilon Y_t) = \mathcal{O}(\pi_t) + \varepsilon E[\mathcal{D}(\pi_t)' Y_t] + O(\varepsilon^2) \quad (3.81)$$

and $\mathcal{D}(\pi)$ is the Frechet derivative. To this end, we compute

$$\begin{aligned}
E[\mu(s_t)'(\pi(s_t, s_{t-}) + \varepsilon Y_t)] &= E[\mu(s_t)'\pi(s_t, s_{t-})] + \varepsilon E[\mu(s_t)'Y_t] \\
E[(\pi(s_t, s_{t-}) + \varepsilon Y_t)'(\Sigma + \Lambda + \bar{\Lambda}_t)(\pi(s_t, s_{t-}) + \varepsilon Y_t)] \\
&= E[\pi(s_t, s_{t-})'(\Sigma + \Lambda + \bar{\Lambda}_t)\pi(s_t, s_{t-})] \\
&+ 2\varepsilon E[\pi(s_t, s_{t-})'(\Sigma + \Lambda + \bar{\Lambda}_t)Y_t] + O(\varepsilon^2) \\
E[(\pi(s_t, s_{t-}) + Y_t)'g_{t+1}\Lambda(\pi(s_{t+1}, s_{t+1-}) + Y_{t+1})] \\
&= E[\pi(s_t, s_{t-})'g_{t+1}\Lambda\pi(s_{t+1}, s_{t+1-})] \\
&+ \varepsilon \left(E[Y_t' E_t[g_{t+1}\Lambda\pi(s_{t+1}, s_{t+1-})]] + E[\pi(s_t, s_{t-})'g_{t+1}\Lambda Y_{t+1}] \right)
\end{aligned} \tag{3.82}$$

Furthermore, by stationarity,

$$E[\pi(s_t, s_{t-})Y_{t+1}] = E[\pi(s_{t-1}, s_{t-1-})Y_t].$$

We conclude that the Frechet derivative is given by

$$\begin{aligned}
\mathcal{D}(\pi) &= -2\mu(s_t) + 2(\Sigma + \Lambda + \bar{\Lambda})\pi(s_t, s_{t-}) \\
&- 2E_t[g_{t+1}\Lambda\pi(s_{t+1}, s_{t+1-})] - 2\Lambda g_t \pi(s_{t-1}, s_{t-1-}),
\end{aligned} \tag{3.83}$$

where $\bar{\Lambda}(s_t) = E_t[\Lambda(s_{t+1})]$, implying that a bounded π is optimal if it satisfies the *integral equation*

$$\pi(s_t, s_{t-}) = (\Sigma + \Lambda + \bar{\Lambda}_t)^{-1} \left(\mu(s_t) + E_t[g_{t+1}\Lambda\pi(s_{t+1}, s_{t+1-})] + \Lambda g_t \pi(s_{t-1}, s_{t-1-}) \right). \tag{3.84}$$

Substituting the Ansatz

$$\pi_t = Q_t + m_t g_t \pi_{t-1} \tag{3.85}$$

into this equation, we get

$$\begin{aligned}
&Q_t + m_t g_t \pi_{t-1} \\
&= (\Sigma + \Lambda + \bar{\Lambda}_t)^{-1} \left(\mu(s_t) \right. \\
&+ \left. E_t[g_{t+1}\Lambda(Q_{t+1} + m_{t+1}g_{t+1}Q_t + m_{t+1}g_{t+1}m_t g_t \pi_{t-1})] + \Lambda g_t \pi_{t-1} \right).
\end{aligned} \tag{3.86}$$

Equating the coefficients on π_{t-1} gives an integral equation for m_t :

$$m_t g_t = (\Sigma + \Lambda + \bar{\Lambda}_t)^{-1} \left(E_t[g_{t+1} \Lambda m_{t+1} g_{t+1} m_t g_t] + \Lambda g_t \right), \quad (3.87)$$

and (3.86) turns into an integral equation for Q_t :

$$Q_t = (\Sigma + \Lambda + \bar{\Lambda}_t)^{-1} \left(\mu(s_t) + E_t[g_{t+1} \Lambda (Q_{t+1} + m_{t+1} g_{t+1} Q_t)] \right). \quad (3.88)$$

Dividing by g_t , we get that (3.87) turns into the required equation (3.69). Furthermore, we can rewrite (3.87) as

$$(I - (\Sigma + \Lambda + \bar{\Lambda}_t)^{-1} E_t[g_{t+1} \Lambda m_{t+1} g_{t+1}]) m_t = (\Sigma + \Lambda + \bar{\Lambda}_t)^{-1} \Lambda, \quad (3.89)$$

which implies

$$(I - (\Sigma + \Lambda + \bar{\Lambda}_t)^{-1} E_t[g_{t+1} \Lambda m_{t+1} g_{t+1}]) = (\Sigma + \Lambda + \bar{\Lambda}_t)^{-1} \Lambda m_t^{-1}. \quad (3.90)$$

After a few algebraic transformations, we get that (3.88) is equivalent to

$$(I - (\Sigma + \Lambda + \bar{\Lambda}_t)^{-1} E_t[g_{t+1} \Lambda m_{t+1} g_{t+1}]) Q_t = (\Sigma + \Lambda + \bar{\Lambda}_t)^{-1} \left(\mu(s_t) + E_t[g_{t+1} \Lambda Q_{t+1}] \right). \quad (3.91)$$

Substituting from (3.90), we can rewrite (3.91) as

$$Q_t = m_t \Lambda^{-1} \left(\mu(s_t) + E_t[g_{t+1} \Lambda Q_{t+1}] \right). \quad (3.92)$$

Defining

$$c_t = M_t \Lambda_t^{-1} \Sigma_t, \quad (3.93)$$

we get

$$Q_t = c_t \text{Markowitz}_t + E_t[m_t \Lambda^{-1} g_{t+1} \Lambda Q_{t+1}]. \quad (3.94)$$

Iterating this equation, we get the required up to the convergence statement. Convergence in L_2 follows directly from the made assumptions. Indeed,

$$E[E_t[X]^2] \leq E[X^2], \quad (3.95)$$

and hence we can ignore $E_t[\cdot]$ when proving convergence. Furthermore, by the made

uniform boundedness assumptions and the uniform positive-definiteness of Σ_t , we have

$$\|q_{t+\tau}\| = \|c_{t+\tau} \text{Markowitz}_{t+\tau}\| \leq K$$

for some $K > 0$, almost surely. Since

$$\|Q_t\| = \|E_t \left[\sum_{\tau=0}^{\infty} N_{t,t+\tau} c_{t+\tau} \text{Markowitz}_{t+\tau} \right]\| \leq \sum_{\tau} \|N_{t,t+\tau} q_{t+\tau}\|, \quad (3.96)$$

to prove the convergence of Q_t it suffices to show that

$$\sum_{\tau} E[q'_{t+\tau} N'_{t,t+\tau} N_{t,t+\tau} q_{t+\tau}]^{1/2} < \infty \quad (3.97)$$

which follows from the made assumptions. Convergence of the series representation for π_t also follows from the made assumptions. \blacksquare

Recall that g_t is the diagonal matrix of $\text{vec}(g_t)$ on the diagonal. For the case when μ is constant, we have that

$$G_t = E_t[\text{vac}(g_{t+1})\text{vec}(g_{t+1})'] = (1 + g_t^w)^{-2} (\Sigma + (1 + r_t^f + \mu_t)(1 + r_t^f + \mu_t)'). \quad (3.98)$$

When g_t^w , r_t^f , and μ_t are all constant, we get that $G_t = G$ is also constant, and we arrive at the following result, which is a direct consequence of Lemma 7.

Lemma 8 *For simplicity, we normalize $\gamma/w = 1$. Suppose that $\mu_t = \mu$, g_t^w, r_t^f are constant. Let $\bar{\Lambda} = \Lambda \circ G$. For any solution m to*

$$m = (\Sigma + \Lambda + \bar{\Lambda})^{-1} \left(((\Lambda m) \circ G)m + \Lambda \right), \quad (3.99)$$

define

$$N_{t,t+\tau} = \prod_{\tau=1}^{\theta} m g_{t-\tau+1} \quad (3.100)$$

and

$$\tilde{N}_{t,t+\tau} = \prod_{\tau=1}^{\theta} m \Lambda^{-1} g_{t-\tau+1} \Lambda. \quad (3.101)$$

Suppose that

$$\sum_{\tau=1}^{\infty} \|E_t[N'_{t,t+\tau} N_{t,t+\tau}]\|^{1/2} < \infty \quad (3.102)$$

and

$$\sum_{\tau=1}^{\infty} \|E_t[\tilde{N}'_{t,t+\tau}\tilde{N}_{t,t+\tau}]\|^{1/2} < \infty. \quad (3.103)$$

Define

$$c = m\Lambda^{-1}\Sigma \quad (3.104)$$

and

$$Q_t = (I - m\Lambda^{-1}\bar{g}\Lambda)^{-1}c\text{Markowitz} \quad (3.105)$$

and let

$$\pi_t = \pi(s_t, s_{t-}) = \sum_{\theta=0}^{\infty} N_{t-\theta,t}Q(s_{t-\theta}). \quad (3.106)$$

Then, all series converge in L_2 , and π_t is optimal among all bounded stationary processes π_t . Furthermore, it satisfies the recursive relationship

$$\pi(s_t, s_{t-}) = Q(s_t) + mg_t\pi(s_{t-1}, s_{t-1-}). \quad (3.107)$$

Properties the Discount Factor m

This section shows some useful properties of the discount factor m solving (3.99). We start with the following observation

Lemma 9 *For simplicity, we normalize $\gamma/w = 1$. A matrix-valued function $m(s_t) = m_t$ solves*

$$m_t = (\Sigma + \Lambda + \bar{\Lambda}_t)^{-1} \left(E_t[g_{t+1}\Lambda m_{t+1}g_{t+1}]m_t + \Lambda \right) \quad (3.108)$$

if and only if $\tilde{m}_t = \Lambda^{1/2}m_t\Lambda^{-1/2}$ solves

$$\tilde{m}_t = \left(\Lambda^{-1/2}\Sigma\Lambda^{-1/2} + I + \Lambda^{-1/2}E_t[g_{t+1}\Lambda^{1/2}(I - \tilde{m}_{t+1})\Lambda^{1/2}g_{t+1}]\Lambda^{-1/2} \right)^{-1}. \quad (3.109)$$

In particular, if G_t is constant, then a matrix m solves

$$m = (\Sigma + \Lambda + \bar{\Lambda})^{-1} \left(((\Lambda m) \circ G)m + \Lambda \right) \quad (3.110)$$

with $\bar{\Lambda} = \Lambda \circ G$ if and only if the matrix $\tilde{m} = \Lambda^{1/2}m\Lambda^{-1/2}$ solves

$$\tilde{m} = \left(\Lambda^{-1/2}\Sigma\Lambda^{-1/2} + I + \Lambda^{-1/2}(G \circ (\Lambda^{1/2}(I - \tilde{m})\Lambda^{1/2}))\Lambda^{-1/2} \right)^{-1}. \quad (3.111)$$

Proof of [. Proof of Lemma 9] We have

$$(\Sigma + \Lambda + \bar{\Lambda}_t)m_t = \left(E_t[g_{t+1}\Lambda m_{t+1}g_{t+1}]m_t + \Lambda \right). \quad (3.112)$$

Writing $m_t = \Lambda^{-1/2}\tilde{m}_t\Lambda^{1/2}$, we get

$$(\Sigma + \Lambda + \bar{\Lambda}_t)\Lambda^{-1/2}\tilde{m}_t\Lambda^{1/2} = \left(E_t[g_{t+1}\Lambda\Lambda^{-1/2}\tilde{m}_{t+1}\Lambda^{1/2}g_{t+1}]\Lambda^{-1/2}\tilde{m}_t\Lambda^{1/2} + \Lambda \right). \quad (3.113)$$

Multiplying by $\Lambda^{-1/2}\tilde{m}_t^{-1}$ from the right and by $\Lambda^{-1/2}$ from the left, we get

$$\Lambda^{-1/2}(\Sigma + \Lambda + \bar{\Lambda}_t)\Lambda^{-1/2} = \left(\Lambda^{-1/2}E_t[g_{t+1}\Lambda^{1/2}\tilde{m}_{t+1}\Lambda^{1/2}g_{t+1}]\Lambda^{-1/2} + \tilde{m}_t^{-1} \right). \quad (3.114)$$

This is equivalent to

$$\tilde{m}_t^{-1} = \Lambda^{-1/2}(\Sigma + \Lambda + \bar{\Lambda}_t - E_t[g_{t+1}\Lambda^{1/2}\tilde{m}_{t+1}\Lambda^{1/2}g_{t+1}])\Lambda^{-1/2}, \quad (3.115)$$

which is, in turn, equivalent to

$$\tilde{m}_t = \left(\Lambda^{-1/2}\Sigma\Lambda^{-1/2} + I + \Lambda^{-1/2}E_t[g_{t+1}\Lambda^{1/2}(I - \tilde{m}_{t+1})\Lambda^{1/2}g_{t+1}]\Lambda^{-1/2} \right)^{-1}. \quad (3.116)$$

In the case when $m_t = m$ is constant and $G_t = G$ is constant, we get

$$E_t[g_{t+1}\Lambda^{1/2}(I - \tilde{m})\Lambda^{1/2}g_{t+1}] = G \circ (\Lambda^{1/2}(I - \tilde{m})\Lambda^{1/2}),$$

and we get the required. ■

Recall that $\mathcal{S}(0, 1)$ is the set of symmetric, positive semi-definite matrices with eigenvalues in $(0, 1)$.

Proposition 18 *Suppose that there exists a solution $\tilde{m} \in \mathcal{S}(0, 1)$ to (3.111). Let $q_* < 1$ be the largest eigenvalue of \tilde{m} . Let $m = \Lambda^{-1/2}\tilde{m}\Lambda^{1/2}$ and*

$$\Pi_{t-\theta, t} = \left(\prod_{\tau=1}^{\theta} m g_{t-\tau+1} \right). \quad (3.117)$$

Then,

$$\lim_{\theta \rightarrow \infty} q_*^{-\theta} E[\|\Pi_{t-\theta, t}(\nu)\|^2] = 0. \quad (3.118)$$

Similarly, if we define

$$\hat{\Pi}_{t,t+\theta} = \prod_{\tau=1}^{\theta} (m\Lambda^{-1}g_{t+\tau-1}\Lambda).$$

Then,

$$\lim_{\theta \rightarrow \infty} q_*^{-\theta} E[\|\hat{\Pi}_{t,t+\theta}\|^2] = 0. \quad (3.119)$$

Proof of [. Proof of Proposition 18] For simplicity, we normalize $\gamma/w = 1$. Recall that \mathcal{S} is the set of symmetric matrices, and $\mathcal{S}(a, b)$ is the set of positive semi-definite matrices with eigenvalues between a and b .

Equation (3.111) can be rewritten as

$$\tilde{m} \left(\Lambda^{-1/2} \Sigma \Lambda^{-1/2} + I + \Lambda^{-1/2} E_t[g_{t+1} \Lambda^{1/2} (I - \tilde{m}_{t+1}) \Lambda^{1/2} g_{t+1}] \Lambda^{-1/2} \right) \tilde{m} = \tilde{m}. \quad (3.120)$$

Define the map

$$\Xi(Z) = \tilde{m} \Lambda^{-1/2} E_t[g_{t+1} \Lambda^{1/2} Z \Lambda^{1/2} g_{t+1}] \Lambda^{-1/2} \tilde{m} \quad (3.121)$$

mapping the cone of positive semi-definite matrices into itself. Then, (3.120) implies that

$$\Xi(I - \tilde{m}) = \tilde{m} - \tilde{m}^2 - \tilde{m} \Lambda^{-1/2} \Sigma \Lambda^{-1/2} \tilde{m} < \tilde{m}(I - \tilde{m}) \leq q_*(I - \tilde{m}). \quad (3.122)$$

Now, the map Ξ leaves the proper cone $\mathcal{S}(0, +\infty)$ invariant, and hence, by the Krein and Rutman (1950) theorem, its spectral radius corresponds to a strictly positive eigenvalue $\lambda_* > 0$. Let $Z \in \mathcal{S}(0, +\infty)$ be the corresponding eigenvector. Then,

$$\Xi(Z) = \lambda_* Z.$$

Since $I - \tilde{m}$ is strictly positive definite, there exists a constant $a_* > 0$ such that $aZ \leq I - \tilde{m}$ if and only if $a \leq a_*$. Then,

$$\lambda a_* Z = \Xi(a_* Z) \leq \Xi(I - \tilde{m}) < q_*(I - \tilde{m})$$

Thus, $(\lambda a_*/q_*)Z \leq I - \tilde{m}$ implying that, by the definition of a_* , we must have $\lambda/q_* < 1$.

Note also that the transformation

$$\hat{\Xi}(Z) = A^{-1} \tilde{m} E_t[\Lambda^{-1/2} g_{t+1} \Lambda^{1/2} A Z A' \Lambda^{1/2} g_{t+1} \Lambda^{-1/2}] \tilde{m} (A')^{-1}$$

is similar to Ξ for any invertible matrix A . Hence, Ξ and $\hat{\Xi}$ have the same spectral radius. Pick $A = \tilde{m} \Lambda^{-1/2}$. Then,

$$\hat{\Xi}(Z) = E_t[g_{t+1} \Lambda^{1/2} \tilde{m} \Lambda^{-1/2} Z \Lambda^{-1/2} \tilde{m} \Lambda^{1/2} g_{t+1}] = E_t[g_{t+1} m' Z m g_{t+1}].$$

By direct calculation,

$$E[\Pi'_{t-\theta,t}\Pi_{t-\theta,t}] = \hat{\Xi}^\theta(I),$$

and

$$E[\hat{\Pi}_{t,t+\theta}\hat{\Pi}'_{t,t+\theta}] = \Xi^\theta(I),$$

and the claim follows. \blacksquare

Proofs of Propositions 12 and 13

Lemma 10 *For simplicity, we normalize $\gamma/w = 1$. Consider the map F mapping the convex set of $\mathcal{S}(0,1)$ -valued matrix functions into itself and defined via*

$$F(\tilde{m}_t) = \left(\Lambda^{-1/2}\Sigma\Lambda^{-1/2} + I + \Lambda^{-1/2}E_t[g_{t+1}\Lambda^{1/2}(I - \tilde{m}_{t+1})\Lambda^{1/2}g_{t+1}]\Lambda^{-1/2} \right)^{-1}. \quad (3.123)$$

This map is strictly monotone increasing in the positive semi-definite order and hence has at least one fixed point in $\mathcal{S}(0,1)$. The set of fixed points has a unique maximal and a unique minimal element. The minimal element is obtained by iterating F on 0. The maximal element is obtained by iterating F on I .

Proof of [Proof of Lemma 10] The proof follows directly from the fact that the map $A \rightarrow A^{-1}$ is monotone decreasing in the positive semi-definite order, and the same is true for the map $\tilde{m}_{t+1} \rightarrow E_t[g_{t+1}\Lambda^{1/2}(I - \tilde{m}_{t+1})\Lambda^{1/2}g_{t+1}]$. \blacksquare

When μ is stochastic, things are a bit more tricky, as shown by the following lemma.

Lemma 11 *Suppose that $\mu(s_t) = \epsilon\tilde{\mu}(s_t)$, $g_t^w = g^w + O(\epsilon)$, $r_t^f = r^f + O(\epsilon)$. Then,*

$$G_t = G + O(\epsilon) \quad (3.124)$$

where $G = E[\text{vec}(g_t)\text{vec}(g_t)']$ and hence, for every solution $\tilde{m} \in \mathcal{S}(0,1)$ to (3.111) and any sufficiently small $\epsilon > 0$ there exists a unique solution \tilde{m}_t to (3.109) satisfying

$$\tilde{m}_t = \tilde{m} + O(\epsilon). \quad (3.125)$$

Proof of [Proof of Lemma 11] The proof follows directly from the implicit function theorem and the fact that the map F from Lemma 10 is strictly monotone on $\mathcal{S}(0,1)$ and (by direct calculation) has a non-degenerate Jacobian. \blacksquare

Proof of [Proof of Proposition 12] By Lemma 10, there exists a $\tilde{m} \in \mathcal{S}(0,1)$ solving (3.111). By Proposition 18, the technical conditions (3.102) and (3.103) are satisfied. Lemma 9 implies that m solves (3.99) and hence Lemma 8 implies that the policy π is optimal. Its uniqueness follows from the strict concavity of the objective. \blacksquare

Proof of [. Proof of Proposition 13] By Lemma 10, there exists a $\tilde{m} \in \mathcal{S}(0, 1)$ solving (3.111). By Lemma 11, there exists a \tilde{m}_t solving (3.109) satisfying $m_t = \tilde{m} + O(\varepsilon)$. By a small modification of the proof of Lemma 18, the technical conditions (3.72) and (3.73) are satisfied. Lemma 9 implies that m_t solves (3.108) and hence Lemma 7 implies that the policy

$$\pi_t^* = \pi(s_t, s_{t-}) = \sum_{\theta=0}^{\infty} N_{t-\theta, t} Q(s_{t-\theta}). \quad (3.126)$$

is optimal with

$$Q_t = E_t \left[\sum_{\tau=0}^{\infty} \tilde{N}_{t, t+\tau} c_{t+\tau} \text{Markowitz}_{t+\tau} \right] \quad (3.127)$$

and

$$N_{t, t+\tau} = \prod_{\tau=1}^{\theta} m_{t-\tau+1} g_{t-\tau+1} \quad (3.128)$$

and

$$\tilde{N}_{t, t+\tau} = \prod_{\tau=1}^{\theta} m_{t-\tau+1} \Lambda^{-1} g_{t-\tau+1} \Lambda. \quad (3.129)$$

and $c_t = m_t \Lambda^{-1} \Sigma$. Its uniqueness follows from the strict concavity of the objective. Now, substituting $m_t = m + O(\varepsilon)$ into these equations, we get that π_t^* differs from (3.20) by $O(\varepsilon)$ (technical conditions (3.72)-(3.73) ensure that the infinite sum also is $O(\varepsilon)$.) The proof is complete. \blacksquare

Proofs of Propositions 14 and 15

We have $m = \Lambda^{-1/2} \tilde{m} \Lambda^{1/2}$ and hence m and \tilde{m} have identical eigenvalues, and the claim of Proposition 14 follows from Lemma 6 and the monotonicity of the map F .

The convergence of m to zero when $w \rightarrow 0$ follows directly from (3.111). When $w \rightarrow \infty$, to prove convergence, we need to show that the technical conditions (3.102) and (3.103) hold uniformly when $w \rightarrow \infty$. By Proposition 18, to this end we need to show that q_* , the maximal eigenvalue of \tilde{m} stays uniformly bounded away from 1. This follows from Lemma 6: Since $\tilde{m} \leq \tilde{m}(\bar{g})$, it suffices to establish this fact for $\tilde{m}(\bar{g})$. From the proof of Lemma 5, we have

$$\tilde{m}(\xi) = \xi^{-1/2} \hat{m} \xi^{-1/2}$$

where

$$\hat{m} = 0.5(\hat{\Sigma} - (\hat{\Sigma}^2 - 4I)^{1/2}) = 2(\hat{\Sigma} + (\hat{\Sigma}^2 - 4I)^{1/2})^{-1} \quad (3.130)$$

and

$$\hat{\Sigma} = \xi^{-1/2} \tilde{\Sigma} \xi^{-1/2} = \gamma w^{-1} \xi^{-1/2} \Lambda^{-1/2} \Sigma \Lambda^{-1/2} \xi^{-1/2} + \xi^{-1} + \xi.$$

and hence $\hat{\Sigma} \rightarrow \xi^{-1} + \xi$ when $w \rightarrow \infty$. Thus,

$$\tilde{m} \rightarrow 0.5\xi^{-1}(\xi^{-1} + \xi - |\xi - \xi^{-1}|)$$

Suppose that $\xi > 1$. Then, we get $\tilde{m} = \xi^{-2} < 1$. If $\xi < 1$, then we get $\tilde{m} \rightarrow 1$.

On the Optimality of Portfolio-ML

Given a function $A(\cdot)$, denote

$$\Pi_t^*(A(\cdot), \vec{s}) = \sum_{\theta=0}^{\infty} \left(\prod_{\tau=1}^{\theta} m g_{t-\tau+1} \right) (I - m)A(s_{t-\theta}). \quad (3.131)$$

We formalize the above observations in the following lemma.

Lemma 12 *Suppose that $\mu(s_t) = \varepsilon \tilde{\mu}(s_t)$ where ε is a small number. Then, the solution to the aim optimization problem*

$$\begin{aligned} \max_{A(\cdot)} E & \left[\mu(s_t)' \Pi_t^*(A(\cdot), \vec{s}) \right. \\ & - \frac{w}{2} \left(\Pi_t^*(A(\cdot), \vec{s}_t) - g_t \Pi_t^*(A(\cdot), \vec{s}_{t-1}) \right)' \Lambda \left(\Pi_t^*(A(\cdot), \vec{s}_t) - g_t \Pi_t^*(A(\cdot), \vec{s}_{t-1}) \right) \\ & \left. - \frac{\gamma}{2} \left(\Pi_t^*(A(\cdot), \vec{s}_t) \right)' \Sigma \left(\Pi_t^*(A(\cdot), \vec{s}_t) \right) \right] \quad (3.132) \end{aligned}$$

coincides with (3.21) up to terms of the order ε^2 .

Lemma 12 reduces the infeasible (due to infinite history dependence) portfolio optimization problem (3.9) to a feasible aim optimization problem (3.132), where only the function $A(s_t)$ of the current state needs to be optimized. As we now show, it is possible to use machine learning methods to further reduce (3.132) to a linear-quadratic problem that can be solved analytically. We will need the following assumption.

Assumption 1 *Let s_{-i} denote the vector of signals for all stocks except i . There exists a function $a(s_i, s_{-i})$ such that $(A(s_t))_i = a_i(s_{i,t}, s_{-i,t})$.*

Assumption 1 is not restrictive and naturally holds whenever $\Lambda_{i,j}$ and $\Sigma_{i,j}$ only depend on s_i, s_j for any pair of stocks i, j . It ensures that the dependence of the aim on signals is the same for all stocks.

Assumption 2 Suppose that a family of functions $\{f_k(s)\}_{k \geq 1}$ has the universal approximation property: For any $\epsilon > 0$ there exists a $P > 0$ and a vector $\beta \in \mathbb{R}^P$ such that

$$\|a(s) - \sum_k \beta_k f_k(s)\|_2 \leq \epsilon. \quad (3.133)$$

Let now $F(s) = (f_k(s_i, s_{-i}))_{i,k=1}^{n,P}$. Then, Assumptions 1 and 2 imply the existence of a vector β such that

$$\|A(s) - F(s)\beta\|_2 \leq \epsilon, \quad (3.134)$$

and hence we can rewrite (3.131) as

$$\pi_t = X_t \beta + O(\epsilon), \quad (3.135)$$

where

$$X_t \equiv \left[\sum_{\theta=0}^{\infty} \left(\prod_{\tau=1}^{\theta} m g_{t-\tau+1} \right) (I - m) f(s_{t-\theta}) \right]. \quad (3.136)$$

Using this formulation for π_t in the objective (3.132), we can rewrite it as

$$\begin{aligned} & E \left[r'_{t+1} X_t \beta - \frac{\gamma}{2} \beta' X_t' \Sigma X_t \beta - \frac{w}{2} (X_t \beta - g_t X_{t-1} \beta)' \Lambda (X_t \beta - g_t X_{t-1} \beta) \right] \\ &= \frac{1}{T} E \left[\underbrace{r'_{t+1} X_t \beta}_{\equiv \tilde{r}'_{t+1}} - \frac{1}{2} \beta' \underbrace{[\gamma X_t' \Sigma X_t + w (X_t - g_t X_{t-1})' \Lambda (X_t - g_t X_{t-1})]}_{\equiv \tilde{\Sigma}_t} \beta \right] \\ &\equiv E[\tilde{r}'_{t+1}] \beta - \frac{1}{2} \beta' E[\tilde{\Sigma}_t] \beta. \end{aligned} \quad (3.137)$$

and the optimal β is given by

$$\beta^* = E[\tilde{\Sigma}_t]^{-1} E[\tilde{r}_{t+1}] \quad (3.138)$$

Uniform boundedness of all coefficients implies that the solution to the approximate optimization problem achieves approximate optimum in (3.132). Thus, we can maximize utility by maximizing this quadratic equation in the unknown parameter vector β .

Proposition 19 (Portfolio-ML) Suppose that $\mu(s_t) = \varepsilon \tilde{\mu}(s_t)$ where ε is a small number. Let β_T be a finite sample counter-part of (3.138). Then, in the limit as $T \rightarrow \infty$, β_T converges to β^* . Furthermore, the optimal portfolio (3.24) achieves the optimal utility (3.9) up to an error of the order $\varepsilon^2 + \epsilon$, where ϵ is defined in (3.133).

Proofs related to Economic Feature Importance

The optimal portfolio admits a simpler analytical expression when transaction costs are small as seen in the following result.

Proposition 20 *Suppose that $\|\Lambda\|$ is small and Λ is diagonal. Then, the aim portfolio A_t is given by*

$$\begin{aligned} A_t &= \text{Markowitz}_t \\ &+ \Sigma^{-1} E_t[\Lambda g_{t+1} \text{Markowitz}_{t+1} - g_{t+1} \Lambda g_{t+1} \text{Markowitz}_t] + O(\|\Lambda\|^2), \end{aligned} \quad (3.139)$$

while investor's utility is given by the same expression as in Proposition 17.

Proof of [. Proof of Proposition 17 and Proposition 20] Under the made assumption, we have

$$m_t = \Sigma^{-1} \Lambda - \Sigma^{-1} (\Lambda + E_t[g_{t+1} \Lambda g_{t+1}]) \Sigma^{-1} \Lambda + O(\varepsilon^3)$$

and therefore

$$c_t = m_t \Lambda^{-1} \Sigma_t. = I - \Sigma^{-1} (\Lambda + E_t[g_{t+1} \Lambda g_{t+1}]) + O(\varepsilon^2).$$

Let $\nu_t = \text{Markowitz}_t$. Then,

$$\begin{aligned} A_t &= (I - m_t)^{-1} \sum_{\tau=0}^{\infty} E_t [M_{t,t+\tau} c_{t+\tau} \text{Markowitz}_{t+\tau}] \\ &= (I - m_t)^{-1} (c_t \nu_t + E_t[m_t \Lambda^{-1} \Lambda g_{t+1} c_{t+1} \nu_{t+1}]) + O(\varepsilon^2) \\ &= (I + \Sigma^{-1} \Lambda) \left(c_t \nu_t + E_t[\Sigma^{-1} \Lambda \Lambda^{-1} \Lambda g_{t+1} c_{t+1} \nu_{t+1}] \right) + O(\varepsilon^2) \\ &= (I + \Sigma^{-1} \Lambda) \left(\left(I - \Sigma^{-1} (\Lambda + E_t[g_{t+1} \Lambda g_{t+1}]) \right) \nu_t \right. \\ &\quad \left. + E_t[\Sigma^{-1} \Lambda \Lambda^{-1} \Lambda g_{t+1} \left(I - \Sigma^{-1} (\Lambda + E_t[g_{t+1} \Lambda g_{t+1}]) \right) \nu_{t+1}] \right) + O(\varepsilon^2) \\ &= (I - \Sigma^{-1} E_t[g_{t+1} \Lambda g_{t+1}]) \nu_t + E_t[\Sigma^{-1} \Lambda g_{t+1} \nu_{t+1}] + O(\varepsilon^2) \\ &= \nu_t + \Sigma^{-1} E_t[\Lambda g_{t+1} \nu_{t+1} - g_{t+1} \Lambda g_{t+1} \nu_t] + O(\varepsilon^2). \end{aligned} \quad (3.140)$$

and

$$\begin{aligned} \pi_t &= m_t g_t \pi_{t-1} + (I - m_t) A_t \\ &= m_t g_t \nu_{t-1} + (I - m_t) (\nu_t + \Sigma^{-1} E_t[\Lambda g_{t+1} (\nu_{t+1} - g_{t+1} \nu_t)]) + O(\varepsilon^2) \\ &= \nu_t + \Sigma^{-1} E_t[\Lambda g_{t+1} (\nu_{t+1} - g_{t+1} \nu_t)] + \Sigma^{-1} \Lambda (g_t \nu_{t-1} - \nu_t) + O(\varepsilon^2) \\ &= \nu_t + \xi_t + O(\varepsilon^2). \end{aligned} \quad (3.141)$$

Hence

$$\begin{aligned}
& E[\mu(x_t)' \pi_t - 0.5 \pi_t' \Sigma \pi_t - 0.5 (\pi_t - g_t \pi_{t-1})' \Lambda (\pi_t - \pi_{t-1})] \\
&= E[\mu(x_t)' \pi_t - 0.5 \pi_t' (\Sigma + \Lambda + E_t[g_{t+1} \Lambda g_{t+1}]) \pi_t + \pi_{t-1}' g_t \Lambda \pi_t] \\
&= E[\mu_t' (\nu_t + \xi_t) \\
&\quad - 0.5 (\nu_t + \xi_t)' (\Sigma + \Lambda + E_t[g_{t+1} \Lambda g_{t+1}]) (\nu_t + \xi_t) \\
&\quad + (\nu_{t-1} + \xi_{t-1})' g_t \Lambda (\nu_t + \xi_t)] \\
&= 0.5 E[\mu_t' \Sigma^{-1} \mu_t] + E[\mu_t' (\Sigma^{-1} E_t[\Lambda g_{t+1} (\nu_{t+1} - g_{t+1} \nu_t)] + \Sigma^{-1} \Lambda (g_t \nu_{t-1} - \nu_t))] \\
&\quad - E[\nu_t' \Sigma \xi_t] - 0.5 E[\nu_t' (\Lambda + E_t[g_{t+1} \Lambda g_{t+1}]) \nu_t] \\
&\quad + E[\nu_{t-1}' g_t \Lambda \nu_{t-1}] \\
&= 0.5 E[\mu_t' \Sigma^{-1} \mu_t] - 0.5 E[\nu_t' (\Lambda + E_t[g_{t+1} \Lambda g_{t+1}]) \nu_t] + E[\nu_t' g_{t+1} \Lambda \nu_{t+1}] \\
&= 0.5 E[\mu_t' \Sigma^{-1} \mu_t] - 0.5 E[(\nu_t - g_t \nu_{t-1})' \Lambda (\nu_t - g_t \nu_{t-1})] + O(\varepsilon^2).
\end{aligned}$$

■

7.3 Data and Empirical Results

Information of Stock Characteristics (Features)

Table AI shows the security characteristics that we use as features for all portfolio methods. The features are a subset of the 153 characteristics used in Jensen et al. (2022b) plus 1-year trailing volatility (rvol_252d),¹⁹ where the subset is chosen to have sufficient coverage in the early parts of our sample. Specifically, we select all features with a non-missing value for at least 70% of our sample by the end of 1952. The cluster assignments in Table AI are also from Jensen et al. (2022b) except rvol_252d, which we assign to the low-risk cluster.

Portfolio Tuning

Panel A shows the optimal hyper-parameters for the RF method that predicts expected returns. The 1 month model is used by all methods except Portfolio-ML, while the expected return over 2-6 and 7-12 months is only used by Multiperiod-ML and Multiperiod-ML*. Panel B shows the optimal hyper-parameters for Portfolio-ML and the second layer of portfolio tuning used by Multiperiod-ML* and Static-ML*. Section 4.3 describes how we choose hyper-parameters and table I the range of possible hyper-parameters.

¹⁹We add 1-year trailing volatility because of its close connection to the covariance matrix. For example, suppose volatility is unrelated to expected returns. In that case, the optimal portfolio should have a higher allocation to low volatility assets, all else equal.

Table AI. Feature Information

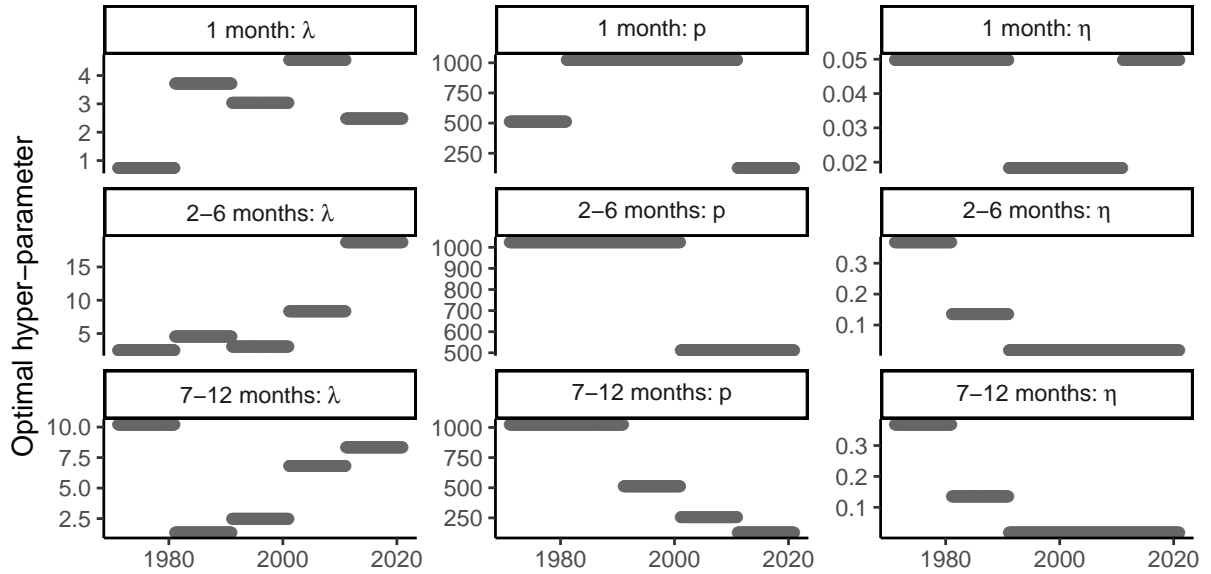
	Characteristic	Cluster		Characteristic	Cluster		Characteristic	Cluster
1	cowc_gr1a	accruals	40	ivol_capm_252d	low_risk	79	opex_at	quality
2	oaccruals_at	accruals	41	ivol_ff3_21d	low_risk	80	qmj_prof	quality
3	oaccruals_ni	accruals	42	rmax1_21d	low_risk	81	qmj_safety	quality
4	taccruals_at	accruals	43	rmax5_21d	low_risk	82	sale_bev	quality
5	taccruals_ni	accruals	44	rvol_21d	low_risk	83	corr_1260d	seasonality
6	fnl_gr1a	debt_issuance	45	turnover_126d	low_risk	84	coskew_21d	seasonality
7	ncol_gr1a	debt_issuance	46	zero_trades_126d	low_risk	85	dbnetis_at	seasonality
8	nfna_gr1a	debt_issuance	47	zero_trades_21d	low_risk	86	kz_index	seasonality
9	noa_at	debt_issuance	48	zero_trades_252d	low_risk	87	lti_gr1a	seasonality
10	aliq_at	investment	49	rvol_252d	low_risk	88	pi_nix	seasonality
11	at_gr1	investment	50	prc_highprc_252d	momentum	89	seas_11_15an	seasonality
12	be_gr1a	investment	51	ret_12_1	momentum	90	seas_11_15na	seasonality
13	capx_gr1	investment	52	ret_3_1	momentum	91	seas_2_5an	seasonality
14	coa_gr1a	investment	53	ret_6_1	momentum	92	seas_6_10an	seasonality
15	col_gr1a	investment	54	ret_9_1	momentum	93	ami_126d	size
16	emp_gr1	investment	55	seas_1_1na	momentum	94	dolvol_126d	size
17	inv_gr1	investment	56	ocf_at_chg1	profit_growth	95	market_equity	size
18	inv_gr1a	investment	57	ret_12_7	profit_growth	96	prc	size
19	lnoa_gr1a	investment	58	sale_emp_gr1	profit_growth	97	iskew_capm_21d	short-term reversal
20	mispricing_mgmt	investment	59	seas_1_1an	profit_growth	98	iskew_ff3_21d	short-term reversal
21	ncoa_gr1a	investment	60	tax_gr1a	profit_growth	99	ret_1_0	short-term reversal
22	nncoa_gr1a	investment	61	dolvol_var_126d	profitability	100	rmax5_rvol_21d	short-term reversal
23	noa_gr1a	investment	62	ebit_bev	profitability	101	rskew_21d	short-term reversal
24	ppeinv_gr1a	investment	63	ebit_sale	profitability	102	at_me	value
25	ret_60_12	investment	64	intrinsic_value	profitability	103	be_me	value
26	sale_gr1	investment	65	ni_be	profitability	104	bev_mev	value
27	seas_2_5na	investment	66	o_score	profitability	105	chcsho_12m	value
28	age	leverage	67	ocf_at	profitability	106	debt_me	value
29	aliq_mat	leverage	68	ope_be	profitability	107	div12m_me	value
30	at_be	leverage	69	ope_bell	profitability	108	ebitda_mev	value
31	bidaskhl_21d	leverage	70	turnover_var_126d	profitability	109	eq_dur	value
32	cash_at	leverage	71	at_turnover	quality	110	eqnpo_12m	value
33	netdebt_me	leverage	72	cop_at	quality	111	fcf_me	value
34	tangibility	leverage	73	cop_atl1	quality	112	ni_me	value
35	beta_60m	low_risk	74	gp_at	quality	113	ocf_me	value
36	beta_dimson_21d	low_risk	75	gp_atl1	quality	114	sale_me	value
37	betatab_1260d	low_risk	76	mispricing_perf	quality	115	seas_6_10na	value
38	betadown_252d	low_risk	77	op_at	quality			
39	ivol_capm_21d	low_risk	78	op_atl1	quality			

Note: The table shows the security characteristics we use as features for the portfolio methods. The characteristics are from Jensen et al. (2022b), and we refer to this paper for details about the construction methodology.

Feature Persistence and Importance across Return Horizons

Figure A2 shows the monthly autocorrelation of all prediction features. Features grouped into themes following Jensen et al. (2022b). We see that most features are highly persistent from month to month but that we also include a substantial amount of fast-moving predictors. These high-frequency predictors are particularly present in the low-risk, seasonality, and short-term reversal themes. Figure A3 shows a measure of feature importance for each of the three models that predicts future returns. The short-term model that predicts returns one month ahead is used by all portfolio methods except Portfolio-ML, while only Multiperiod-ML uses the two other longer-term models. Notably, feature importance for the short-term model differs from the others by distributing importance more evenly across themes. In contrast, the two longer-term models mainly use value and momentum features.

Panel A: Expected Return Tuning



Panel B: Portfolio Tuning

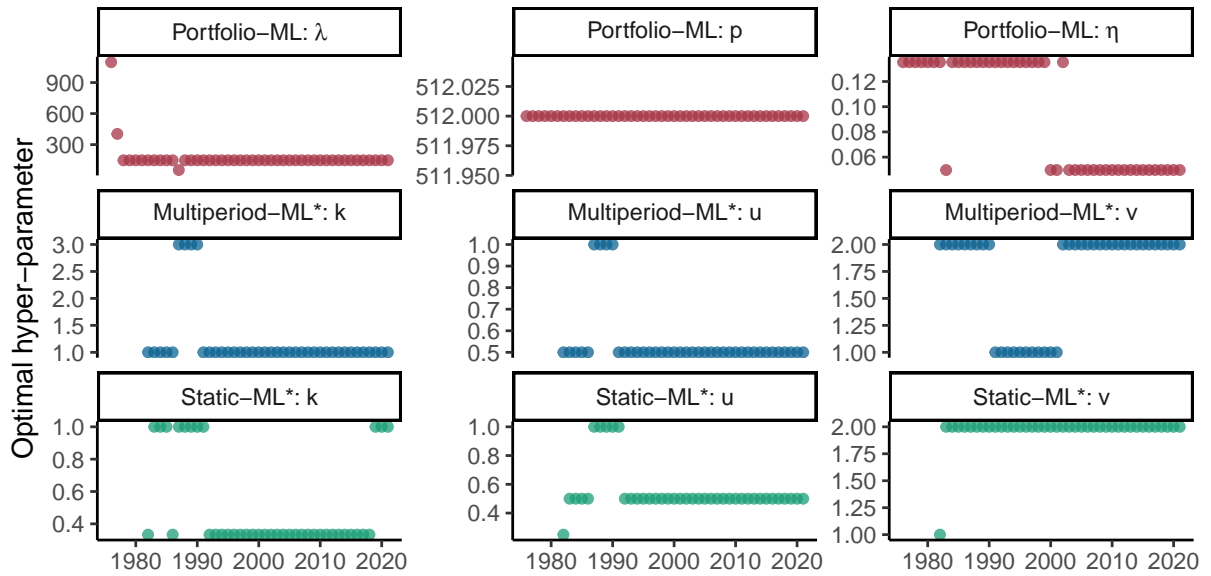


Figure A1. Optimal Portfolio Hyper-parameters

Note: Panel A shows the optimal hyper-parameters used for predicting expected returns via. Ridge regression of RF transformed features. Panel B shows the optimal hyper-parameters for selecting portfolio weights for Portfolio-ML, Multiperiod-ML*, and Static-ML*. We show the range of possible hyper-parameters in table I.

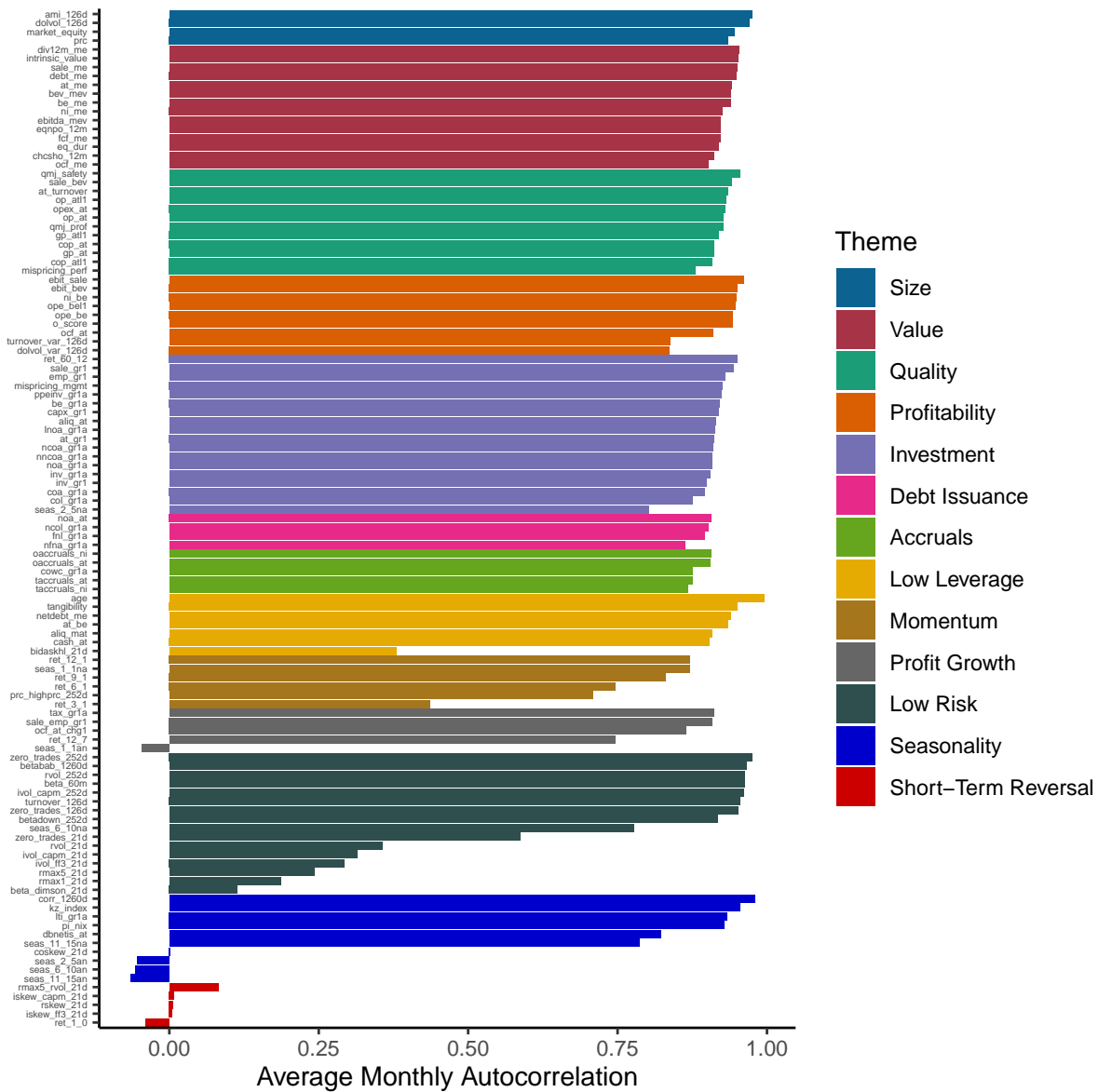


Figure A2. Feature Autocorrelation

Note: The figure shows the monthly autocorrelation for each feature in our sample. We first compute each feature’s monthly autocorrelation for all stocks with at least five years of monthly data. Next, we average the stock-level autocorrelations to arrive at the final estimate. The features are grouped by theme and sorted by average theme autocorrelation.

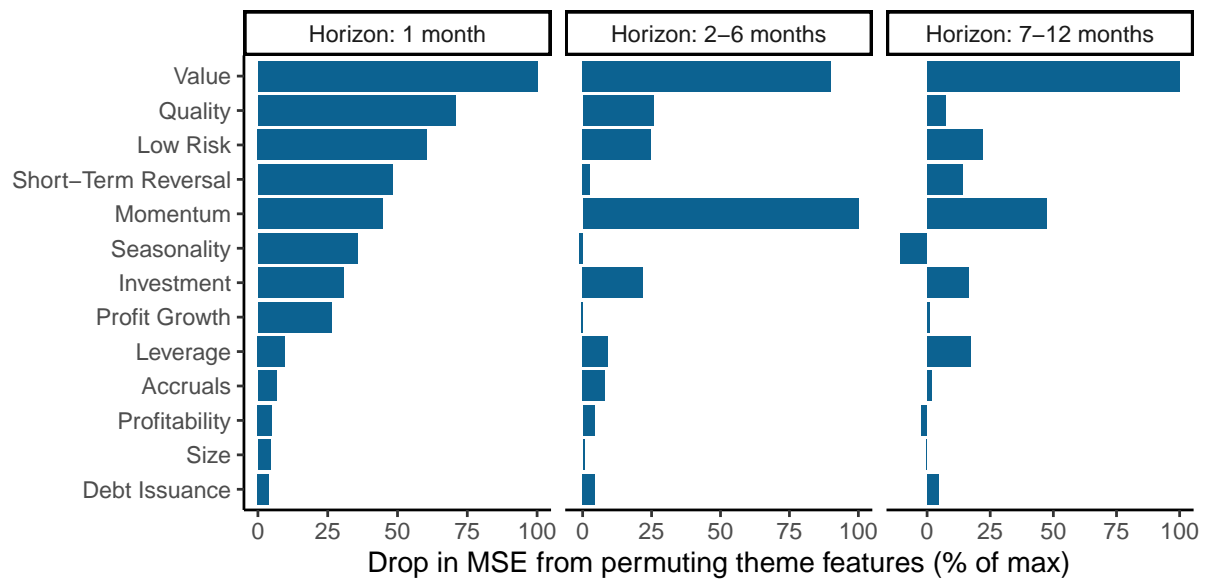


Figure A3. Feature Importance across Return Horizons

Note: The figure shows feature importance for the three random feature-based models that predict returns in month $t + 1$, the average return over month $t + 2$ to $t + 6$, and the average return over month $t + 7$ to $t + 12$. For each model, we randomly permute the associated features for each theme while keeping all other features at their actual value. We then implement each method based on this counterfactual data and measure feature importance as the difference in the mean-squared error relative to the implementation that uses the actual data. For comparability, we re-scale the difference by scaling all differences by the largest difference. Hence, feature importance is measured relative to the best feature theme within a specific horizon.

Bibliography

- Acharya, V. and L. H. Pedersen (2005). Asset pricing with liquidity risk. *Journal of Financial Economics* 77, 375–410.
- Amihud, Y. and H. Mendelson (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics* 17, 223–249.
- Amromin, G. and S. A. Sharpe (2014). From the horse’s mouth: Economic conditions and investor expectations of risk and return. *Management Science* 60(4), 845–866.
- Asness, C. and A. Frazzini (2013). The devil in HML’s details. *The Journal of Portfolio Management* 39(4), 49–68.
- Asness, C., T. Moskowitz, and L. H. Pedersen (2013). Value and momentum everywhere. *The Journal of Finance* 68(3), 929–985.
- Avramov, D., S. Cheng, and L. Metzker (2021). Machine learning versus economic restrictions: Evidence from stock return predictability. *Management Science*, forthcoming.
- Balduzzi, P. and A. W. Lynch (1999). Transaction costs and predictability: some utility cost calculations. *Journal of Financial Economics* 52(1), 47–78.
- Bali, T. G., H. Beckmeyer, M. Moerke, and F. Weigert (2021). Option return predictability with machine learning and big data. *Available at SSRN 3895984*.
- Bali, T. G., A. Goyal, D. Huang, F. Jiang, and Q. Wen (2022). Predicting corporate bond returns: Merton meets machine learning. *Georgetown McDonough School of Business Research Paper* (3686164), 20–110.
- Barberis, N. and M. Huang (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review* 98(5), 2066–2100.
- Barberis, N., L. J. Jin, and B. Wang (2021). Prospect Theory and Stock Market Anomalies. *Journal of Finance* (0), 1–49.
- Barberis, N., A. Shleifer, and R. Vishny (1998). A model of investor sentiment. *Journal of financial economics* 49(3), 307–343.

- Barillas, F. and J. Shanken (2018). Comparing asset pricing models. *The Journal of Finance* 73(2), 715–754.
- Bekaert, G., E. C. Engstrom, and N. R. Xu (2022). The time variation in risk appetite and uncertainty. *Management Science* 68(6), 3975–4004.
- Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* 57(1), 289–300.
- Benjamini, Y. and D. Yekutieli (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* 29(4), 1165–1188.
- Berk, J. B., R. C. Green, and V. Naik (1999). Optimal investment, growth options, and security returns. *Journal of Finance* 54(5), 1553–1607.
- Berk, J. B. and J. H. Van Binsbergen (2016). Assessing asset pricing models using revealed preference. *Journal of Financial Economics* 119(1), 1–23.
- Berry, D. A. and Y. Hochberg (1999). Bayesian perspectives on multiple comparisons. *Journal of Statistical Planning and Inference* 82(1-2), 215–227.
- Bettis, R. A. (2012). The search for asterisks: Compromised statistical tests and flawed theories. *Strategic Management Journal* 33(1), 108–113.
- Bharath, S. T. and T. Shumway (2008). Forecasting default with the merton distance to default model. *The Review of Financial Studies* 21(3), 1339–1369.
- Binsbergen, J. H. V. and C. C. Opp (2019). Real anomalies. *The Journal of finance* 74(4), 1659–1706.
- Black, F. (1973). Yes, Virginia, There Is Hope: Tests of the Value Line Ranking System. *Financial Analysts Journal* 29(5), 10–92.
- Bordalo, P., N. Gennaioli, R. L. Porta, and A. Shleifer (2019). Diagnostic expectations and stock returns. *The Journal of Finance* 74(6), 2839–2874.
- Bouchaud, J.-P., P. Krueger, A. Landier, and D. Thesmar (2019). Sticky expectations and the profitability anomaly. *The Journal of Finance* 74(2), 639–674.
- Brandt, M. W., P. Santa-Clara, and R. Valkanov (2009). Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *The Review of Financial Studies* 22(9), 3411–3447.
- Brav, A., R. Lehavy, and R. Michaely (2005). Using expectations to test asset pricing models. *Financial Management* 34(3), 31–64.

- Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics* 7(3), 265–296.
- Breiman, L. (2001). Random forests. *Machine learning* 45(1), 5–32.
- Britten-Jones, M. (1999). The sampling error in estimates of mean-variance efficient portfolio weights. *The Journal of Finance* 54(2), 655–671.
- Bryzgalova, S., J. Huang, and C. Julliard (2019). Bayesian solutions for the factor zoo: We just ran two quadrillion models. *Available at SSRN*.
- Cakici, N. and A. Zaremba (2022). Empirical asset pricing via machine learning: The global edition. *Available at SSRN 4028525*.
- Campbell, J. (2017). *Financial Decisions and Markets: A Course in Asset Pricing* (1st ed.). Princeton University Press.
- Campbell, J. Y. and T. Vuolteenaho (2004). Bad beta, good beta. *American Economic Review* 94(5), 1249–1275.
- Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *the Journal of Finance* LII(1), 57–82.
- Carlson, M., A. Fisher, and R. Giammarino (2004). Corporate investment and asset price dynamics: Implications for the cross-section of returns. *The Journal of Finance* 59(6), 2577–2603.
- Cassella, S., B. Golez, H. Gulen, and P. Kelly (2021). Horizon bias and the term structure of equity returns. *Available at SSRN 3328970*.
- Cassella, S., B. Golez, H. Gulen, and P. Kelly (2022). Motivated beliefs in macroeconomic expectations. *Available at SSRN 3328970*.
- Chen, A. Y. (2020). The limits of p-hacking: Some thought experiments. *The Journal of Finance, forthcoming*.
- Chen, A. Y. and M. Velikov (2021). Zeroing in on the expected returns of anomalies. *Working paper, Federal Reserve Board*.
- Chen, A. Y. and T. Zimmermann (2020a). Open source cross-sectional asset pricing. *Working paper, Board of Governors of the Federal Reserve System*.
- Chen, A. Y. and T. Zimmermann (2020b). Publication bias and the cross-section of stock returns. *The Review of Asset Pricing Studies* 10(2), 249–289.
- Chen, L., M. Pelger, and J. Zhu (2021). Deep learning in asset pricing. *Available at SSRN 3350138*.

- Chen, T. and C. Guestrin (2016). XGBoost: A Scalable Tree Boosting System. *KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.
- Chinco, A., A. Neuhierl, and M. Weber (2020). Estimating the anomaly base rate. *Journal of Financial Economics*.
- Choi, D., W. Jiang, and C. Zhang (2021). Alpha go everywhere: Machine learning and international stock returns. *Available at SSRN 3489679*.
- Chordia, T., A. Goyal, and A. Saretto (2020). Anomalies and false rejections. *The Review of Financial Studies* 33(5), 2134–2179.
- CNBC (1998). Berkshire hathaway annual meeting - 1998. <https://buffett.cnbc.com/video/1998/05/04/morning-session---1998-berkshire-hathaway-annual-meeting.html?%5C&start=3518%5C&end=3716>. Accessed: 2022-04-13.
- Cochrane, J. H. (2005). *Asset pricing: (Revised Edition)*. Princeton University Press.
- Cochrane, J. H. (2011). Presidential address: Discount rates. *Journal of Finance* 66, 1047–1108.
- Collin-Dufresne, P., K. Daniel, and M. Sağlam (2020). Liquidity regimes and optimal dynamic asset allocation. *Journal of Financial Economics* 136(2), 379–406.
- Constantinides, G. M. (1986). Capital market equilibrium with transaction costs. *Journal of Political Economy* 94, 842–862.
- Dahlquist, M. and M. Ibert (2021). Return expectations and portfolios: Evidence from large asset managers. *Swedish House of Finance Research Paper* (21-1).
- Daniel, K., D. Hirshleifer, and A. Subrahmanyam (1998). Investor psychology and security market under-and overreactions. *the Journal of Finance* 53(6), 1839–1885.
- Daniel, K., D. Hirshleifer, and L. Sun (2020). Short-and long-horizon behavioral factors. *The review of financial studies* 33(4), 1673–1736.
- Davis, M. H. and A. R. Norman (1990). Portfolio selection with transaction costs. *Mathematics of operations research* 15(4), 676–713.
- De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990a). Noise trader risk in financial markets. *Journal of political Economy* 98(4), 703–738.
- De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990b). Positive feedback investment strategies and destabilizing rational speculation. *the Journal of Finance* 45(2), 379–395.

- Detzel, A. L., R. Novy-Marx, and M. Velikov (2021). Model selection with transaction costs. *Available at SSRN 3805379*.
- Efron, B. (2007). Size, power and false discovery rates. *The Annals of Statistics* 35(4), 1351–1377.
- Efron, B. (2012). *Large-scale inference: empirical Bayes methods for estimation, testing, and prediction*, Volume 1. Cambridge University Press.
- Efron, B. and T. Hastie (2016). *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*. Institute of Mathematical Statistics Monographs. Cambridge University Press.
- Efron, B. and R. Tibshirani (2002). Empirical bayes methods and false discovery rates for microarrays. *Genetic epidemiology* 23(1), 70–86.
- Elton, E. J., M. J. Gruber, and J. Spitzer (2006). Improved estimates of correlation coefficients and their impact on optimum portfolios. *European Financial Management* 12(3), 303–318.
- Engelberg, J., R. D. McLean, and J. Pontiff (2018). Anomalies and news. *The Journal of Finance* 73(5), 1971–2001.
- Engelberg, J., R. D. McLean, and J. Pontiff (2020). Analysts and anomalies. *Journal of Accounting and Economics* 69(1), 101249.
- Engle, R. and B. Kelly (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics* 30(2), 212–228.
- Fama, E. F. and K. R. French (1993a). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33(1), 3–56.
- Fama, E. F. and K. R. French (1993b). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33(1), 3–56.
- Fama, E. F. and K. R. French (1993c). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33(1), 3–56.
- Fama, E. F. and K. R. French (2008). Dissecting anomalies. *Journal of Finance* 63(4), 1653–1678.
- Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. *Journal of Financial Economics* 116(1), 1–22.
- Feng, G., S. Giglio, and D. Xiu (2020). Taming the factor zoo: A test of new factors. *The Journal of Finance* 75(3), 1327–1370.
- Frazzini, A., R. Israel, and T. J. Moskowitz (2018). Trading costs. *Available at SSRN 3229719*.

- Frazzini, A. and L. H. Pedersen (2014). Betting against beta. *Journal of Financial Economics* 111(1), 1 – 25.
- Freyberger, J., A. Neuhierl, and M. Weber (2020). Dissecting characteristics nonparametrically. *The Review of Financial Studies* 33(5), 2326–2377.
- Gârleanu, N. and L. H. Pedersen (2013). Dynamic trading with predictable returns and transaction costs. *Journal of Finance* forthcoming.
- Gârleanu, N. and L. H. Pedersen (2016). Dynamic portfolio choice with frictions. *Journal of Economic Theory* 165, 487–516.
- Gelman, A. (2016, Aug). Bayesian inference completely solves the multiple comparisons problem.
- Gelman, A., J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin (2013). *Bayesian Data Analysis, Third Edition*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis.
- Gelman, A., J. Hill, and M. Yajima (2012). Why we (usually) don’t have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness* 5(2), 189–211.
- Gibbons, M. R., S. A. Ross, and J. Shanken (1989). A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society*, 1121–1152.
- Gormsen, N. J. and K. Huber (2022). Equity factors and firms’ perceived cost of capital. Technical report, SSRN Working Paper 3712699.
- Green, J., J. R. Hand, and X. F. Zhang (2017). The characteristics that provide independent information about average us monthly stock returns. *The Review of Financial Studies* 30(12), 4389–4436.
- Greenland, S. and A. Hofman (2019). Multiple comparisons controversies are about context and costs, not frequentism versus bayesianism. *European journal of epidemiology* 34(9), 801–808.
- Greenland, S. and J. M. Robins (1991). Empirical-bayes adjustments for multiple comparisons are sometimes useful. *Epidemiology*, 244–251.
- Greenwood, R. and A. Shleifer (2014). Expectations of returns and expected returns. *The Review of Financial Studies* 27(3), 714–746.
- Gu, S., B. Kelly, and D. Xiu (2020a). Empirical Asset Pricing via Machine Learning. *Review of Financial Studies* 33(5), 2223–2273.
- Gu, S., B. Kelly, and D. Xiu (2020b). Empirical asset pricing via machine learning. *The Review of Financial Studies* 33(5), 2223–2273.
- Gu, S., B. Kelly, and D. Xiu (2021). Autoencoder asset pricing models. *Journal of Econometrics* 222(1), 429–450.

- Hamermesh, D. S. (2007). Replication in economics. *Canadian Journal of Economics/Revue canadienne d'économique* 40(3), 715–733.
- Han, Y., A. He, D. Rapach, and G. Zhou (2021). Expected stock returns and firm characteristics: E-lasso, assessment, and implications. *Assessment, and Implications (September 10, 2021)*.
- Hansen, L. P. and R. Jagannathan (1991). Implications of security market data for models of dynamic economies. *Journal of political economy* 99(2), 225–262.
- Harvey, C. R. (2017). Presidential address: The scientific outlook in financial economics. *The Journal of Finance* 72(4), 1399–1440.
- Harvey, C. R., Y. Liu, and H. Zhu (2016a). ...and the Cross-Section of Expected Returns. *Review of Financial Studies* 29(1), 5–68.
- Harvey, C. R., Y. Liu, and H. Zhu (2016b). ...and the cross-section of expected returns. *The Review of Financial Studies* 29(1), 5–68.
- Heston, S. L. and R. Sadka (2008). Seasonality in the cross-section of stock returns. *Journal of Financial Economics* 87(2), 418–445.
- Hong, H. and D. A. Sraer (2016). Speculative betas. *The Journal of Finance* 71(5), 2095–2144.
- Hong, H. and J. C. Stein (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of finance* 54(6), 2143–2184.
- Hou, K., C. Xue, and L. Zhang (2015). Digesting anomalies: An investment approach. *Review of Financial Studies* 28(3), 650–705.
- Hou, K., C. Xue, and L. Zhang (2020a). Replicating Anomalies. *Review of Financial Studies* 33(5), 2019–2133.
- Hou, K., C. Xue, and L. Zhang (2020b). Replicating anomalies. *The Review of Financial Studies* 33(5), 2019–2133.
- Ilmanen, A., R. Israel, T. J. Moskowitz, R. Lee, and A. K. Thapar (2021). How do factor premia vary over time? a century of evidence. *Journal of Investment Management, Forthcoming*.
- Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS medicine* 2(8), e124.
- Jacobs, H. and S. Muller (2020). Anomalies across the globe: Once public, no longer existent? *Journal of Financial Economics* 135(1), 213–230.
- Jensen, T. I., B. T. Kelly, S. Malamud, and L. H. Pedersen (2022). Machine learning and the implementable efficient frontier. *Available at SSRN 4187217*.

- Jensen, T. I., B. T. Kelly, and L. H. Pedersen (2022a). Is there a replication crisis in finance? *Journal of Finance*, forthcoming.
- Jensen, T. I., B. T. Kelly, and L. H. Pedersen (2022b). Is there a replication crisis in finance? *Journal of Finance*, forthcoming.
- Jo, C., C. Lin, and Y. You (2022). Subjective risk-return trade-off. Available at SSRN 4096443.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Kelly, B. T., S. Malamud, and K. Zhou (2022). The virtue of complexity in return prediction.
- Kelly, B. T., D. Palhares, and S. Pruitt (2022). Modeling corporate bond returns. *Journal of Finance*.
- Kelly, B. T., S. Pruitt, and Y. Su (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics* 134(3), 501–524.
- Kogan, L. and D. Papanikolaou (2013). Firm characteristics and stock returns: The role of investment-specific shocks. *Review of Financial Studies* 26(11), 2718–2759.
- Kogan, L. and D. Papanikolaou (2014). Growth Opportunities, Technology Shocks, and Asset Prices. *Journal of Finance* LXIX(2).
- Koijen, R. S., T. J. Moskowitz, L. H. Pedersen, and E. B. Vrugt (2018). Carry. *Journal of Financial Economics* 127(2), 197–225.
- Koijen, R. S. and M. Yogo (2019). A demand system approach to asset pricing. *Journal of Political Economy* 127(4), 1475–1515.
- Kothari, S. (2001). Capital markets research in accounting. *Journal of accounting and economics* 31(1-3), 105–231.
- Kozak, S., S. Nagel, and S. Santosh (2020). Shrinking the cross-section. *Journal of Financial Economics* 135(2), 271–292.
- Krein, M. G. and M. Rutman (1950). *Linear operators leaving invariant a cone in a Banach space*. Number 26. American Mathematical Society.
- La Porta, R. (1996). Expectations and the Cross-Section of Stock Returns. *The Journal of Finance* 51(5), 1715–1742.
- Leippold, M., Q. Wang, and W. Zhou (2022). Machine learning in the chinese stock market. *Journal of Financial Economics* 145(2), 64–82.
- Lettau, M. and J. A. Wachter (2007). Why is long-horizon equity less risky? a duration-based explanation of the value premium. *The Journal of Finance* 62(1), 55–92.

- Li, S. A., V. DeMiguel, and A. Martin-Utrera (2020). Which factors with price-impact costs? *Victor and Martin-Utrera, Alberto, Which Factors with Price-Impact Costs*.
- Linnainmaa, J. T. and M. R. Roberts (2018). The history of the cross-section of stock returns. *The Review of Financial Studies* 31(7), 2606–2649.
- Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics* 47(1), 13–37.
- Lochstoer, L. A. and T. Muir (2022). Volatility expectations and returns. *Journal of Finance*, forthcoming.
- Lui, D., S. Markov, and A. Tamayo (2007). What makes a stock risky? Evidence from sell-side analysts' risk ratings. *Journal of Accounting Research* 45(3), 629–665.
- Lynch, A. W. and P. Balduzzi (2000). Predictability and transaction costs: the impact on rebalancing rules and behavior. *Journal of Finance* 55, 2285–2309.
- Lynch, P. and J. Rothchild (2000). *One up on Wall Street: how to use what you already know to make money in the market*. Simon and Schuster.
- Maniadis, Z., F. Tufano, and J. A. List (2017). To replicate or not to replicate? exploring reproducibility in economics through the lens of a model and a pilot study. *The Economic Journal* 127, F209–F235.
- Maritz, J. S. (2018). *Empirical Bayes methods with applications*. CRC Press.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance* 7(1), 77–91.
- Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. Yale University Press.
- Martin, I. W. and C. Wagner (2019). What is the expected return on a stock? *The Journal of Finance* 74(4), 1887–1929.
- McLean, R. D. and J. Pontiff (2016). Does academic research destroy stock return predictability? *The Journal of Finance* 71(1), 5–32.
- Mehra, R. and E. C. Prescott (1985). The equity premium: A puzzle. *Journal of monetary Economics* 15(2), 145–161.
- Molnar, C. (2022). *Interpretable Machine Learning* (2 ed.).
- Moskowitz, T. J., Y. H. Ooi, and L. H. Pedersen (2012). Time series momentum. *Journal of financial economics* 104(2), 228–250.
- Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica* 34(4), 768–783.

- MSCI Barra (2007). Barra Risk Model Handbook.
- Murtagh, F. and P. Legendre (2014). Ward’s hierarchical agglomerative clustering method: which algorithms implement ward’s criterion? *Journal of classification* 31(3), 274–295.
- Nadarajah, S. and S. Kotz (2008). Exact distribution of the max/min of two gaussian random variables. *IEEE Transactions on very large scale integration (VLSI) systems* 16(2), 210–212.
- Nagel, S. and Z. Xu (2022a). Asset pricing with fading memory. *The Review of Financial Studies* 35(5), 2190–2245.
- Nagel, S. and Z. Xu (2022b). Dynamics of subjective risk premia. Technical report, National Bureau of Economic Research.
- Newey, W. K. and K. D. West (1987, May). A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55(3), 703–708.
- Nosek, B. A., J. R. Spies, and M. Motyl (2012). Scientific utopia: Ii. restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science* 7(6), 615–631.
- Pedersen, L. H. (2021). Game On : Social Networks and Markets.
- Rahimi, A. and B. Recht (2007). Random features for large-scale kernel machines. *Advances in neural information processing systems* 20.
- Sharot, T. (2011). The optimism bias. *Current biology* 21(23), R941–R945.
- Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance* 19(3), 425–442.
- Shefrin, H. and M. Statman (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of finance* 40(3), 777–790.
- Shumway, T. (1997). The delisting bias in crsp data. *The Journal of Finance* 52(1), 327–340.
- Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review* 71(3), 289–315.
- Stambaugh, R. F. and Y. Yuan (2017). Mispricing factors. *The Review of Financial Studies* 30(4), 1270–1315.
- Sutherland, D. J. and J. Schneider (2015). On the error of random fourier features. *arXiv preprint arXiv:1506.02785*.
- ValueLine (2021). *Value Line Database Manual*.

- Vissing-Jorgensen, A. (2003). Perspectives on behavioral finance: Does “irrationality” disappear with wealth? evidence from expectations and actions. *NBER macroeconomics annual* 18, 139–194.
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58(301), 236–244.
- Welch, I. (2019). Reproducing, extending, updating, replicating, reexamining, and reconciling. *Critical Finance Review* 8(1-2), 301–304.
- Zhang, L. (2005). The value premium. *The Journal of Finance* 60(1), 67–103.

TITLER I PH.D.SERIEN:

– a Field Study of the Rise and Fall of a Bottom-Up Process

2004

1. Martin Grieger
Internet-based Electronic Marketplaces and Supply Chain Management
2. Thomas Basbøll
*LIKENESS
A Philosophical Investigation*
3. Morten Knudsen
*Beslutningens vaklen
En systemteoretisk analyse af moderniseringen af et amtskommunalt sundhedsvæsen 1980-2000*
4. Lars Bo Jeppesen
*Organizing Consumer Innovation
A product development strategy that is based on online communities and allows some firms to benefit from a distributed process of innovation by consumers*
5. Barbara Dragsted
*SEGMENTATION IN TRANSLATION AND TRANSLATION MEMORY SYSTEMS
An empirical investigation of cognitive segmentation and effects of integrating a TM system into the translation process*
6. Jeanet Hardis
*Sociale partnerskaber
Et socialkonstruktivistisk casestudie af partnerskabsaktørers virkelighedsopfattelse mellem identitet og legitimitet*
7. Henriette Hallberg Thygesen
System Dynamics in Action
8. Carsten Mejer Plath
Strategisk Økonomistyring
9. Annemette Kjærgaard
Knowledge Management as Internal Corporate Venturing
10. Knut Arne Hovdal
*De professionelle i endring
Norsk ph.d., ej til salg gennem Samfundslitteratur*
11. Søren Jeppesen
*Environmental Practices and Greening Strategies in Small Manufacturing Enterprises in South Africa
– A Critical Realist Approach*
12. Lars Frode Frederiksen
*Industriel forskningsledelse
– på sporet af mønstre og samarbejde i danske forskningsintensive virksomheder*
13. Martin Jes Iversen
*The Governance of GN Great Nordic
– in an age of strategic and structural transitions 1939-1988*
14. Lars Pynt Andersen
*The Rhetorical Strategies of Danish TV Advertising
A study of the first fifteen years with special emphasis on genre and irony*
15. Jakob Rasmussen
Business Perspectives on E-learning
16. Sof Thrane
*The Social and Economic Dynamics of Networks
– a Weberian Analysis of Three Formalised Horizontal Networks*
17. Lene Nielsen
Engaging Personas and Narrative Scenarios – a study on how a user-centered approach influenced the perception of the design process in the e-business group at AstraZeneca
18. S.J Valstad
*Organisationsidentitet
Norsk ph.d., ej til salg gennem Samfundslitteratur*

19. Thomas Lyse Hansen
Six Essays on Pricing and Weather risk in Energy Markets
20. Sabine Madsen
Emerging Methods – An Interpretive Study of ISD Methods in Practice
21. Evis Sinani
The Impact of Foreign Direct Investment on Efficiency, Productivity Growth and Trade: An Empirical Investigation
22. Bent Meier Sørensen
Making Events Work Or, How to Multiply Your Crisis
23. Pernille Schnoor
*Brand Ethos
Om troværdige brand- og virksomhedsidentiteter i et retorisk og diskursteoretisk perspektiv*
24. Sidsel Fabech
*Von welchem Österreich ist hier die Rede?
Diskursive forhandlinger og magtkampe mellem rivaliserende nationale identitetskonstruktioner i østrigske pressediskurser*
25. Klavs Odgaard Christensen
*Sprogpolitik og identitetsdannelse i flersprogede forbundsstater
Et komparativt studie af Schweiz og Canada*
26. Dana B. Minbaeva
Human Resource Practices and Knowledge Transfer in Multinational Corporations
27. Holger Højlund
*Markedets politiske fornuft
Et studie af velfærdens organisering i perioden 1990-2003*
28. Christine Mølgaard Frandsen
*A.s erfaring
Om mellemværendets praktik i en transformation af mennesket og subjektiviteten*
29. Sine Nørholm Just
The Constitution of Meaning – A Meaningful Constitution? Legitimacy, identity, and public opinion in the debate on the future of Europe
- 2005**
1. Claus J. Varnes
Managing product innovation through rules – The role of formal and structured methods in product development
2. Helle Hedegaard Hein
Mellem konflikt og konsensus – Dialogudvikling på hospitalsklinikker
3. Axel Rosenø
Customer Value Driven Product Innovation – A Study of Market Learning in New Product Development
4. Søren Buhl Pedersen
*Making space
An outline of place branding*
5. Camilla Funck Ellehave
*Differences that Matter
An analysis of practices of gender and organizing in contemporary workplaces*
6. Rigmor Madeleine Lond
Styring af kommunale forvaltninger
7. Mette Aagaard Andreassen
Supply Chain versus Supply Chain Benchmarking as a Means to Managing Supply Chains
8. Caroline Aggestam-Pontoppidan
*From an idea to a standard
The UN and the global governance of accountants' competence*
9. Norsk ph.d.
10. Vivienne Heng Ker-ni
An Experimental Field Study on the

- Effectiveness of Grocer Media Advertising*
Measuring Ad Recall and Recognition, Purchase Intentions and Short-Term Sales
11. Allan Mortensen
Essays on the Pricing of Corporate Bonds and Credit Derivatives
12. Remo Stefano Chiari
Figure che fanno conoscere
Itinerario sull'idea del valore cognitivo e espressivo della metafora e di altri tropi da Aristotele e da Vico fino al cognitivismo contemporaneo
13. Anders McIlquham-Schmidt
Strategic Planning and Corporate Performance
An integrative research review and a meta-analysis of the strategic planning and corporate performance literature from 1956 to 2003
14. Jens Geersbro
The TDF – PMI Case
Making Sense of the Dynamics of Business Relationships and Networks
15. Mette Andersen
Corporate Social Responsibility in Global Supply Chains
Understanding the uniqueness of firm behaviour
16. Eva Boxenbaum
Institutional Genesis: Micro – Dynamic Foundations of Institutional Change
17. Peter Lund-Thomsen
Capacity Development, Environmental Justice NGOs, and Governance: The Case of South Africa
18. Signe Jarlov
Konstruktioner af offentlig ledelse
19. Lars Stæhr Jensen
Vocabulary Knowledge and Listening Comprehension in English as a Foreign Language
- An empirical study employing data elicited from Danish EFL learners*
20. Christian Nielsen
Essays on Business Reporting
Production and consumption of strategic information in the market for information
21. Marianne Thejls Fischer
Egos and Ethics of Management Consultants
22. Annie Bekke Kjær
Performance management i Process-innovation
– belyst i et social-konstruktivistisk perspektiv
23. Suzanne Dee Pedersen
GENTAGELSENS METAMORFOSE
Om organisering af den kreative gørem i den kunstneriske arbejdspraksis
24. Benedikte Dorte Rosenbrink
Revenue Management
Økonomiske, konkurrencemæssige & organisatoriske konsekvenser
25. Thomas Riise Johansen
Written Accounts and Verbal Accounts
The Danish Case of Accounting and Accountability to Employees
26. Ann Fogelgren-Pedersen
The Mobile Internet: Pioneering Users' Adoption Decisions
27. Birgitte Rasmussen
Ledelse i fællesskab – de tillidsvalgtes fornyende rolle
28. Gitte Thit Nielsen
Remerger
– skabende ledelseskrafter i fusion og opkøb
29. Carmine Gioia
A MICROECONOMETRIC ANALYSIS OF MERGERS AND ACQUISITIONS

30. Ole Hinz
Den effektive forandringsleder: pilot, pædagog eller politiker?
Et studie i arbejdslederens meningstilskrivninger i forbindelse med vellykket gennemførelse af ledelsesinitierede forandringsprojekter
31. Kjell-Åge Gotvassli
Et praksisbasert perspektiv på dynamiske læringsnettverk i toppidretten
Norsk ph.d., ej til salg gennem Samfundslitteratur
32. Henriette Langstrup Nielsen
Linking Healthcare
An inquiry into the changing performances of web-based technology for asthma monitoring
33. Karin Tweddell Levinsen
Virtuel Uddannelsespraksis
Master i IKT og Læring – et casestudie i hvordan proaktiv proceshåndtering kan forbedre praksis i virtuelle læringsmiljøer
34. Anika Liversage
Finding a Path
Labour Market Life Stories of Immigrant Professionals
35. Kasper Elmquist Jørgensen
Studier i samspillet mellem stat og erhvervsliv i Danmark under 1. verdenskrig
36. Finn Janning
A DIFFERENT STORY
Seduction, Conquest and Discovery
37. Patricia Ann Plackett
Strategic Management of the Radical Innovation Process
Leveraging Social Capital for Market Uncertainty Management
- 2006**
1. Christian Vintergaard
Early Phases of Corporate Venturing
2. Niels Rom-Poulsen
Essays in Computational Finance
3. Tina Brandt Husman
Organisational Capabilities, Competitive Advantage & Project-Based Organisations
The Case of Advertising and Creative Good Production
4. Mette Rosenkrands Johansen
Practice at the top
– how top managers mobilise and use non-financial performance measures
5. Eva Parum
Corporate governance som strategisk kommunikations- og ledelsesværktøj
6. Susan Aagaard Petersen
Culture's Influence on Performance Management: The Case of a Danish Company in China
7. Thomas Nicolai Pedersen
The Discursive Constitution of Organizational Governance – Between unity and differentiation
The Case of the governance of environmental risks by World Bank environmental staff
8. Cynthia Selin
Volatile Visions: Transactions in Anticipatory Knowledge
9. Jesper Banghøj
Financial Accounting Information and Compensation in Danish Companies
10. Mikkel Lucas Overby
Strategic Alliances in Emerging High-Tech Markets: What's the Difference and does it Matter?
11. Tine Aage
External Information Acquisition of Industrial Districts and the Impact of Different Knowledge Creation Dimensions

- A case study of the Fashion and Design Branch of the Industrial District of Montebelluna, NE Italy*
12. Mikkel Flyverbom
*Making the Global Information Society Governable
On the Governmentality of Multi-Stakeholder Networks*
 13. Anette Grønning
*Personen bag
Tilstedevær i e-mail som inter-aktionsform mellem kunde og medarbejder i dansk forsikringskontekst*
 14. Jørn Helder
*One Company – One Language?
The NN-case*
 15. Lars Bjerregaard Mikkelsen
*Differing perceptions of customer value
Development and application of a tool for mapping perceptions of customer value at both ends of customer-supplier dyads in industrial markets*
 16. Lise Granerud
*Exploring Learning
Technological learning within small manufacturers in South Africa*
 17. Esben Rahbek Pedersen
*Between Hopes and Realities:
Reflections on the Promises and Practices of Corporate Social Responsibility (CSR)*
 18. Ramona Samson
*The Cultural Integration Model and European Transformation.
The Case of Romania*
- 2007**
1. Jakob Vestergaard
*Discipline in The Global Economy
Panopticism and the Post-Washington Consensus*
 2. Heidi Lund Hansen
*Spaces for learning and working
A qualitative study of change of work, management, vehicles of power and social practices in open offices*
 3. Sudhanshu Rai
*Exploring the internal dynamics of software development teams during user analysis
A tension enabled Institutionalization Model; "Where process becomes the objective"*
 4. Norsk ph.d.
Ej til salg gennem Samfundslitteratur
 5. Serden Ozcan
*EXPLORING HETEROGENEITY IN ORGANIZATIONAL ACTIONS AND OUTCOMES
A Behavioural Perspective*
 6. Kim Sundtoft Hald
*Inter-organizational Performance Measurement and Management in Action
– An Ethnography on the Construction of Management, Identity and Relationships*
 7. Tobias Lindeberg
*Evaluative Technologies
Quality and the Multiplicity of Performance*
 8. Merete Wedell-Wedellsborg
*Den globale soldat
Identitetsdannelse og identitetsledelse i multinationale militære organisationer*
 9. Lars Frederiksen
*Open Innovation Business Models
Innovation in firm-hosted online user communities and inter-firm project ventures in the music industry
– A collection of essays*
 10. Jonas Gabrielsen
Retorisk toposlære – fra statisk 'sted' til persuasiv aktivitet

11. Christian Moldt-Jørgensen
*Fra meningsløs til meningsfuld evaluering.
Anvendelsen af studentertilfredsheds-målinger på de korte og mellemlange videregående uddannelser set fra et psykodynamisk systemperspektiv*
12. Ping Gao
*Extending the application of actor-network theory
Cases of innovation in the telecommunications industry*
13. Peter Mejlby
*Frihed og fængsel, en del af den samme drøm?
Et phronetisk baseret casestudie af frigørelsens og kontrollens sam-eksistens i værdibaseret ledelse!*
14. Kristina Birch
Statistical Modelling in Marketing
15. Signe Poulsen
*Sense and sensibility:
The language of emotional appeals in insurance marketing*
16. Anders Bjerre Trolle
Essays on derivatives pricing and dynamic asset allocation
17. Peter Feldhütter
Empirical Studies of Bond and Credit Markets
18. Jens Henrik Eggert Christensen
Default and Recovery Risk Modeling and Estimation
19. Maria Theresa Larsen
*Academic Enterprise: A New Mission for Universities or a Contradiction in Terms?
Four papers on the long-term implications of increasing industry involvement and commercialization in academia*
20. Morten Wellendorf
*Postimplementering af teknologi i den offentlige forvaltning
Analyser af en organisations kontinuerlige arbejde med informations-teknologi*
21. Ekaterina Mhaanna
Concept Relations for Terminological Process Analysis
22. Stefan Ring Thorbjørnsen
*Forsvaret i forandring
Et studie i officerers kapabiliteter under påvirkning af omverdenens forandringspres mod øget styring og læring*
23. Christa Breum Amhøj
Det selvskabte medlemskab om managementstaten, dens styringsteknologier og indbyggere
24. Karoline Bromose
*Between Technological Turbulence and Operational Stability
– An empirical case study of corporate venturing in TDC*
25. Susanne Justesen
*Navigating the Paradoxes of Diversity in Innovation Practice
– A Longitudinal study of six very different innovation processes – in practice*
26. Luise Noring Henler
*Conceptualising successful supply chain partnerships
– Viewing supply chain partnerships from an organisational culture perspective*
27. Mark Mau
*Kampen om telefonen
Det danske telefonvæsen under den tyske besættelse 1940-45*
28. Jakob Halskov
The semiautomatic expansion of existing terminological ontologies using knowledge patterns discovered

- on the WWW – an implementation and evaluation*
29. Gergana Koleva
European Policy Instruments Beyond Networks and Structure: The Innovative Medicines Initiative
 30. Christian Geisler Asmussen
Global Strategy and International Diversity: A Double-Edged Sword?
 31. Christina Holm-Petersen
*Stolthed og fordom
Kultur- og identitetsarbejde ved skabelsen af en ny sengeafdeling gennem fusion*
 32. Hans Peter Olsen
*Hybrid Governance of Standardized States
Causes and Contours of the Global Regulation of Government Auditing*
 33. Lars Bøge Sørensen
Risk Management in the Supply Chain
 34. Peter Aagaard
*Det unikkes dynamikker
De institutionelle mulighedsbetingelser bag den individuelle udforskning i professionelt og frivilligt arbejde*
 35. Yun Mi Antorini
*Brand Community Innovation
An Intrinsic Case Study of the Adult Fans of LEGO Community*
 36. Joachim Lynggaard Boll
*Labor Related Corporate Social Performance in Denmark
Organizational and Institutional Perspectives*
- 2008**
1. Frederik Christian Vinten
Essays on Private Equity
 2. Jesper Clement
Visual Influence of Packaging Design on In-Store Buying Decisions
 3. Marius Brostrøm Kousgaard
*Tid til kvalitetsmåling?
– Studier af indrulleringsprocesser i forbindelse med introduktionen af kliniske kvalitetsdatabaser i speciallægepraksissektoren*
 4. Irene Skovgaard Smith
*Management Consulting in Action
Value creation and ambiguity in client-consultant relations*
 5. Anders Rom
*Management accounting and integrated information systems
How to exploit the potential for management accounting of information technology*
 6. Marina Candi
Aesthetic Design as an Element of Service Innovation in New Technology-based Firms
 7. Morten Schnack
*Teknologi og tværfaglighed
– en analyse af diskussionen omkring indførelse af EPJ på en hospitalsafdeling*
 8. Helene Balslev Clausen
Juntos pero no revueltos – un estudio sobre emigrantes norteamericanos en un pueblo mexicano
 9. Lise Justesen
*Kunsten at skrive revisionsrapporter.
En beretning om forvaltningsrevisions beretninger*
 10. Michael E. Hansen
The politics of corporate responsibility: CSR and the governance of child labor and core labor rights in the 1990s
 11. Anne Roepstorff
Holdning for handling – en etnologisk undersøgelse af Virksomheders Sociale Ansvar/CSR

12. Claus Bajlum
Essays on Credit Risk and Credit Derivatives
13. Anders Bojesen
The Performative Power of Competence – an Inquiry into Subjectivity and Social Technologies at Work
14. Satu Reijonen
*Green and Fragile
A Study on Markets and the Natural Environment*
15. Ilduara Busta
*Corporate Governance in Banking
A European Study*
16. Kristian Anders Hvass
*A Boolean Analysis Predicting Industry Change: Innovation, Imitation & Business Models
The Winning Hybrid: A case study of isomorphism in the airline industry*
17. Trine Paludan
*De uvidende og de udviklingsparate
Identitet som mulighed og restriktion blandt fabriksarbejdere på det aftayloriserede fabriksgulv*
18. Kristian Jakobsen
Foreign market entry in transition economies: Entry timing and mode choice
19. Jakob Elming
Syntactic reordering in statistical machine translation
20. Lars Brømsøe Termansen
*Regional Computable General Equilibrium Models for Denmark
Three papers laying the foundation for regional CGE models with agglomeration characteristics*
21. Mia Reinholt
The Motivational Foundations of Knowledge Sharing
22. Frederikke Krogh-Meibom
*The Co-Evolution of Institutions and Technology
– A Neo-Institutional Understanding of Change Processes within the Business Press – the Case Study of Financial Times*
23. Peter D. Ørberg Jensen
OFFSHORING OF ADVANCED AND HIGH-VALUE TECHNICAL SERVICES: ANTECEDENTS, PROCESS DYNAMICS AND FIRMLEVEL IMPACTS
24. Pham Thi Song Hanh
Functional Upgrading, Relational Capability and Export Performance of Vietnamese Wood Furniture Producers
25. Mads Vangkilde
*Why wait?
An Exploration of first-mover advantages among Danish e-grocers through a resource perspective*
26. Hubert Buch-Hansen
*Rethinking the History of European Level Merger Control
A Critical Political Economy Perspective*
- 2009**
1. Vivian Lindhardtsen
From Independent Ratings to Communal Ratings: A Study of CWA Raters' Decision-Making Behaviours
2. Guðrið Weihe
Public-Private Partnerships: Meaning and Practice
3. Chris Nøkkentved
*Enabling Supply Networks with Collaborative Information Infrastructures
An Empirical Investigation of Business Model Innovation in Supplier Relationship Management*
4. Sara Louise Muhr
Wound, Interrupted – On the Vulnerability of Diversity Management

5. Christine Sestoft
Forbrugeradfærd i et Stats- og Livsformsteoretisk perspektiv
6. Michael Pedersen
Tune in, Breakdown, and Reboot: On the production of the stress-fit self-managing employee
7. Salla Lutz
Position and Reposition in Networks – Exemplified by the Transformation of the Danish Pine Furniture Manufacturers
8. Jens Forssbæk
Essays on market discipline in commercial and central banking
9. Tine Murphy
Sense from Silence – A Basis for Organised Action
How do Sensemaking Processes with Minimal Sharing Relate to the Reproduction of Organised Action?
10. Sara Malou Strandvad
Inspirations for a new sociology of art: A sociomaterial study of development processes in the Danish film industry
11. Nicolaas Mouton
On the evolution of social scientific metaphors: A cognitive-historical enquiry into the divergent trajectories of the idea that collective entities – states and societies, cities and corporations – are biological organisms.
12. Lars Andreas Knutsen
Mobile Data Services: Shaping of user engagements
13. Nikolaos Theodoros Korfiatis
Information Exchange and Behavior
A Multi-method Inquiry on Online Communities
14. Jens Albæk
Forestillinger om kvalitet og tværfaglighed på sygehuse
– skabelse af forestillinger i læge- og plejegrupperne angående relevans af nye idéer om kvalitetsudvikling gennem tolkningsprocesser
15. Maja Lotz
The Business of Co-Creation – and the Co-Creation of Business
16. Gitte P. Jakobsen
Narrative Construction of Leader Identity in a Leader Development Program Context
17. Dorte Hermansen
“Living the brand” som en brandorienteret dialogisk praxis: Om udvikling af medarbejdernes brandorienterede dømmekraft
18. Aseem Kinra
Supply Chain (logistics) Environmental Complexity
19. Michael Nørager
How to manage SMEs through the transformation from non innovative to innovative?
20. Kristin Wallevik
Corporate Governance in Family Firms
The Norwegian Maritime Sector
21. Bo Hansen Hansen
Beyond the Process
Enriching Software Process Improvement with Knowledge Management
22. Annemette Skot-Hansen
Franske adjektivisk afledte adverbier, der tager præpositionssyntagmer indledt med præpositionen à som argumenter
En valensgrammatisk undersøgelse
23. Line Gry Knudsen
Collaborative R&D Capabilities
In Search of Micro-Foundations

24. Christian Scheuer
*Employers meet employees
Essays on sorting and globalization*
25. Rasmus Johnsen
*The Great Health of Melancholy
A Study of the Pathologies of Perfor-
mativity*
26. Ha Thi Van Pham
*Internationalization, Competitiveness
Enhancement and Export Performance
of Emerging Market Firms:
Evidence from Vietnam*
27. Henriette Balieu
*Kontrolbegrebets betydning for kausa-
tivalternationen i spansk
En kognitiv-typologisk analyse*
- 2010**
1. Yen Tran
*Organizing Innovation in Turbulent
Fashion Market
Four papers on how fashion firms crea-
te and appropriate innovation value*
2. Anders Raastrup Kristensen
*Metaphysical Labour
Flexibility, Performance and Commit-
ment in Work-Life Management*
3. Margrét Sigrún Sigurdardóttir
*Dependently independent
Co-existence of institutional logics in
the recorded music industry*
4. Ásta Dis Óladóttir
*Internationalization from a small do-
mestic base:
An empirical analysis of Economics and
Management*
5. Christine Secher
*E-deltagelse i praksis – politikernes og
forvaltningens medkonstruktion og
konsekvenserne heraf*
6. Marianne Stang Våland
*What we talk about when we talk
about space:*
7. Rex Degnegaard
*Strategic Change Management
Change Management Challenges in
the Danish Police Reform*
8. Ulrik Schultz Brix
*Værdi i rekruttering – den sikre beslut-
ning
En pragmatisk analyse af perception
og synliggørelse af værdi i rekrutte-
rings- og udvælgelsesarbejdet*
9. Jan Ole Similä
*Kontraktsledelse
Relasjonen mellom virksomhetsledelse
og kontraktshåndtering, belyst via fire
norske virksomheter*
10. Susanne Boch Waldorff
*Emerging Organizations: In between
local translation, institutional logics
and discourse*
11. Brian Kane
*Performance Talk
Next Generation Management of
Organizational Performance*
12. Lars Ohnemus
*Brand Thrust: Strategic Branding and
Shareholder Value
An Empirical Reconciliation of two
Critical Concepts*
13. Jesper Schlamovitz
*Håndtering af usikkerhed i film- og
byggeprojekter*
14. Tommy Moesby-Jensen
*Det faktiske livs forbindtlighed
Førsokratisk informeret, ny-aristotelisk
ἦθος-tænkning hos Martin Heidegger*
15. Christian Fich
*Two Nations Divided by Common
Values
French National Habitus and the
Rejection of American Power*

16. Peter Beyer
Processer, sammenhængskraft og fleksibilitet
Et empirisk casestudie af omstillingsforløb i fire virksomheder
17. Adam Buchhorn
Markets of Good Intentions
Constructing and Organizing Biogas Markets Amid Fragility and Controversy
18. Cecilie K. Moesby-Jensen
Social læring og fælles praksis
Et mixed method studie, der belyser læringskonsekvenser af et lederkursus for et praksisfællesskab af offentlige mellemledere
19. Heidi Boye
Fødevarer og sundhed i sen-modernismen
– En indsigt i hyggefænomenet og de relaterede fødevarerpraksisser
20. Kristine Munkgård Pedersen
Flygtige forbindelser og midlertidige mobiliseringer
Om kulturel produktion på Roskilde Festival
21. Oliver Jacob Weber
Causes of Intercompany Harmony in Business Markets – An Empirical Investigation from a Dyad Perspective
22. Susanne Ekman
Authority and Autonomy
Paradoxes of Modern Knowledge Work
23. Anette Frey Larsen
Kvalitetsledelse på danske hospitaler
– Ledelsernes indflydelse på introduktion og vedligeholdelse af kvalitetsstrategier i det danske sundhedsvæsen
24. Toyoko Sato
Performativity and Discourse: Japanese Advertisements on the Aesthetic Education of Desire
25. Kenneth Brinch Jensen
Identifying the Last Planner System
Lean management in the construction industry
26. Javier Busquets
Orchestrating Network Behavior for Innovation
27. Luke Patey
The Power of Resistance: India's National Oil Company and International Activism in Sudan
28. Mette Vedel
Value Creation in Triadic Business Relationships. Interaction, Interconnection and Position
29. Kristian Tørning
Knowledge Management Systems in Practice – A Work Place Study
30. Qingxin Shi
An Empirical Study of Thinking Aloud Usability Testing from a Cultural Perspective
31. Tanja Juul Christiansen
Corporate blogging: Medarbejderes kommunikative handlekraft
32. Malgorzata Ciesielska
Hybrid Organisations. A study of the Open Source – business setting
33. Jens Dick-Nielsen
Three Essays on Corporate Bond Market Liquidity
34. Sabrina Speiermann
Modstandens Politik
Kampagnestyling i Velfærdsstaten. En diskussion af trafikcampagners styringspotentiale
35. Julie Uldam
Fickle Commitment. Fostering political engagement in 'the flighty world of online activism'

36. Annegrete Juul Nielsen
Traveling technologies and transformations in health care
37. Athur Mühlen-Schulte
Organising Development Power and Organisational Reform in the United Nations Development Programme
38. Louise Rygaard Jonas
Branding på butiksgulvet Et case-studie af kultur- og identitetsarbejdet i Kvickly
- 2011**
1. Stefan Fraenkel
Key Success Factors for Sales Force Readiness during New Product Launch A Study of Product Launches in the Swedish Pharmaceutical Industry
2. Christian Plesner Rossing
International Transfer Pricing in Theory and Practice
3. Tobias Dam Hede
Samtalekunst og ledelsesdisciplin – en analyse af coachingsdiskursens genealogi og governmentality
4. Kim Pettersson
Essays on Audit Quality, Auditor Choice, and Equity Valuation
5. Henrik Merkelsen
The expert-lay controversy in risk research and management. Effects of institutional distances. Studies of risk definitions, perceptions, management and communication
6. Simon S. Torp
Employee Stock Ownership: Effect on Strategic Management and Performance
7. Mie Harder
Internal Antecedents of Management Innovation
8. Ole Helby Petersen
Public-Private Partnerships: Policy and Regulation – With Comparative and Multi-level Case Studies from Denmark and Ireland
9. Morten Krogh Petersen
'Good' Outcomes. Handling Multiplicity in Government Communication
10. Kristian Tangsgaard Hvelplund
Allocation of cognitive resources in translation - an eye-tracking and key-logging study
11. Moshe Yonatany
The Internationalization Process of Digital Service Providers
12. Anne Vestergaard
Distance and Suffering Humanitarian Discourse in the age of Mediatization
13. Thorsten Mikkelsen
Personlighedens indflydelse på forretningsrelationer
14. Jane Thostrup Jagd
Hvorfor fortsætter fusionsbølgen ud-over "the tipping point"? – en empirisk analyse af information og kognitioner om fusioner
15. Gregory Gimpel
Value-driven Adoption and Consumption of Technology: Understanding Technology Decision Making
16. Thomas Stengade Sønderkov
Den nye mulighed Social innovation i en forretningsmæssig kontekst
17. Jeppe Christoffersen
Donor supported strategic alliances in developing countries
18. Vibeke Vad Baunsgaard
Dominant Ideological Modes of Rationality: Cross functional

- integration in the process of product innovation*
19. Throstur Olaf Sigurjonsson
Governance Failure and Iceland's Financial Collapse
 20. Allan Sall Tang Andersen
Essays on the modeling of risks in interest-rate and inflation markets
 21. Heidi Tscherning
Mobile Devices in Social Contexts
 22. Birgitte Gorm Hansen
Adapting in the Knowledge Economy Lateral Strategies for Scientists and Those Who Study Them
 23. Kristina Vaarst Andersen
Optimal Levels of Embeddedness The Contingent Value of Networked Collaboration
 24. Justine Grønbæk Pors
Noisy Management A History of Danish School Governing from 1970-2010
 25. Stefan Linder
Micro-foundations of Strategic Entrepreneurship Essays on Autonomous Strategic Action
 26. Xin Li
Toward an Integrative Framework of National Competitiveness An application to China
 27. Rune Thorbjørn Clausen
Værdifuld arkitektur Et eksplorativt studie af bygningers rolle i virksomheders værdiskabelse
 28. Monica Viken
Markedsundersøkelser som bevis i varemerke- og markedsføringsrett
 29. Christian Wymann
Tattooing The Economic and Artistic Constitution of a Social Phenomenon
 30. Sanne Frandsen
Productive Incoherence A Case Study of Branding and Identity Struggles in a Low-Prestige Organization
 31. Mads Stenbo Nielsen
Essays on Correlation Modelling
 32. Ivan Häuser
Følelse og sprog Etablering af en ekspressiv kategori, eksemplificeret på russisk
 33. Sebastian Schwenen
Security of Supply in Electricity Markets
- 2012**
1. Peter Holm Andreasen
The Dynamics of Procurement Management - A Complexity Approach
 2. Martin Haulrich
Data-Driven Bitext Dependency Parsing and Alignment
 3. Line Kirkegaard
Konsulenten i den anden nat En undersøgelse af det intense arbejdsliv
 4. Tonny Stenheim
Decision usefulness of goodwill under IFRS
 5. Morten Lind Larsen
Produktiviteten, vækst og velfærd Industrirådet og efterkrigstidens Danmark 1945 - 1958
 6. Petter Berg
Cartel Damages and Cost Asymmetries
 7. Lynn Kahle
Experiential Discourse in Marketing A methodical inquiry into practice and theory
 8. Anne Roelsgaard Obling
Management of Emotions in Accelerated Medical Relationships

9. Thomas Frandsen
Managing Modularity of Service Processes Architecture
10. Carina Christine Skovmøller
*CSR som noget særligt
Et casestudie om styring og menings-
skabelse i relation til CSR ud fra en
intern optik*
11. Michael Tell
*Fradragsbeskæring af selskabers
finansieringsudgifter
En skatteretlig analyse af SEL §§ 11,
11B og 11C*
12. Morten Holm
*Customer Profitability Measurement
Models
Their Merits and Sophistication
across Contexts*
13. Katja Joo Dyppel
*Beskatning af derivater
En analyse af dansk skatteret*
14. Esben Anton Schultz
*Essays in Labor Economics
Evidence from Danish Micro Data*
15. Carina Risvig Hansen
*"Contracts not covered, or not fully
covered, by the Public Sector Directive"*
16. Anja Svejgaard Pors
*Iværksættelse af kommunikation
- patientfigurer i hospitalets strategiske
kommunikation*
17. Frans Bévort
*Making sense of management with
logics
An ethnographic study of accountants
who become managers*
18. René Kallestrup
*The Dynamics of Bank and Sovereign
Credit Risk*
19. Brett Crawford
*Revisiting the Phenomenon of Interests
in Organizational Institutionalism
The Case of U.S. Chambers of
Commerce*
20. Mario Daniele Amore
Essays on Empirical Corporate Finance
21. Arne Stjernholm Madsen
*The evolution of innovation strategy
Studied in the context of medical
device activities at the pharmaceutical
company Novo Nordisk A/S in the
period 1980-2008*
22. Jacob Holm Hansen
*Is Social Integration Necessary for
Corporate Branding?
A study of corporate branding
strategies at Novo Nordisk*
23. Stuart Webber
*Corporate Profit Shifting and the
Multinational Enterprise*
24. Helene Ratner
*Promises of Reflexivity
Managing and Researching
Inclusive Schools*
25. Therese Strand
*The Owners and the Power: Insights
from Annual General Meetings*
26. Robert Gavin Strand
*In Praise of Corporate Social
Responsibility Bureaucracy*
27. Nina Sormunen
*Auditor's going-concern reporting
Reporting decision and content of the
report*
28. John Bang Mathiasen
*Learning within a product development
working practice:
- an understanding anchored
in pragmatism*
29. Philip Holst Riis
*Understanding Role-Oriented Enterprise
Systems: From Vendors to Customers*
30. Marie Lisa Dacanay
*Social Enterprises and the Poor
Enhancing Social Entrepreneurship and
Stakeholder Theory*

31. Fumiko Kano Glückstad
Bridging Remote Cultures: Cross-lingual concept mapping based on the information receiver's prior-knowledge
32. Henrik Barslund Fosse
Empirical Essays in International Trade
33. Peter Alexander Albrecht
*Foundational hybridity and its reproduction
Security sector reform in Sierra Leone*
34. Maja Rosenstock
*CSR - hvor svært kan det være?
Kulturanalytisk casestudie om udfordringer og dilemmaer med at forankre Coops CSR-strategi*
35. Jeanette Rasmussen
*Tweens, medier og forbrug
Et studie af 10-12 årige danske børns brug af internettet, opfattelse og forståelse af markedsføring og forbrug*
36. Ib Tunby Gulbrandsen
*'This page is not intended for a US Audience'
A five-act spectacle on online communication, collaboration & organization.*
37. Kasper Aalling Teilmann
Interactive Approaches to Rural Development
38. Mette Mogensen
*The Organization(s) of Well-being and Productivity
(Re)assembling work in the Danish Post*
39. Søren Friis Møller
*From Disinterestedness to Engagement
Towards Relational Leadership In the Cultural Sector*
40. Nico Peter Berhausen
Management Control, Innovation and Strategic Objectives – Interactions and Convergence in Product Development Networks
41. Balder Onarheim
*Creativity under Constraints
Creativity as Balancing 'Constrainedness'*
42. Haoyong Zhou
Essays on Family Firms
43. Elisabeth Naima Mikkelsen
*Making sense of organisational conflict
An empirical study of enacted sense-making in everyday conflict at work*
- 2013**
1. Jacob Lyngsie
Entrepreneurship in an Organizational Context
2. Signe Groth-Brodersen
*Fra ledelse til selvet
En socialpsykologisk analyse af forholdet imellem selvledelse, ledelse og stress i det moderne arbejdsliv*
3. Nis Høyrup Christensen
Shaping Markets: A Neoinstitutional Analysis of the Emerging Organizational Field of Renewable Energy in China
4. Christian Edelvold Berg
*As a matter of size
THE IMPORTANCE OF CRITICAL MASS AND THE CONSEQUENCES OF SCARCITY FOR TELEVISION MARKETS*
5. Christine D. Isakson
*Coworker Influence and Labor Mobility
Essays on Turnover, Entrepreneurship and Location Choice in the Danish Maritime Industry*
6. Niels Joseph Jerne Lennon
*Accounting Qualities in Practice
Rhizomatic stories of representational faithfulness, decision making and control*
7. Shannon O'Donnell
*Making Ensemble Possible
How special groups organize for collaborative creativity in conditions of spatial variability and distance*

8. Robert W. D. Veitch
*Access Decisions in a Partly-Digital World
Comparing Digital Piracy and Legal Modes for Film and Music*
9. Marie Mathiesen
*Making Strategy Work
An Organizational Ethnography*
10. Arisa Shollo
The role of business intelligence in organizational decision-making
11. Mia Kaspersen
The construction of social and environmental reporting
12. Marcus Møller Larsen
The organizational design of offshoring
13. Mette Ohm Rørdam
*EU Law on Food Naming
The prohibition against misleading names in an internal market context*
14. Hans Peter Rasmussen
*GIV EN GED!
Kan giver-idealtyper forklare støtte til velgørenhed og understøtte relationsopbygning?*
15. Ruben Schachtenhaufen
Fonetisk reduktion i dansk
16. Peter Koerver Schmidt
*Dansk CFC-beskatning
I et internationalt og komparativt perspektiv*
17. Morten Froholdt
*Strategi i den offentlige sektor
En kortlægning af styringsmæssig kontekst, strategisk tilgang, samt anvendte redskaber og teknologier for udvalgte danske statslige styrelser*
18. Annette Camilla Sjørup
*Cognitive effort in metaphor translation
An eye-tracking and key-logging study*
19. Tamara Stucchi
*The Internationalization of Emerging Market Firms:
A Context-Specific Study*
20. Thomas Lopdrup-Hjorth
*"Let's Go Outside":
The Value of Co-Creation*
21. Ana Alačovska
*Genre and Autonomy in Cultural Production
The case of travel guidebook production*
22. Marius Gudmand-Høyer
*Stemningssindssygdommenes historie i det 19. århundrede
Omtydningen af melankolien og manien som bipolære stemningslidelser i dansk sammenhæng under hensyn til dannelsen af det moderne følelseslivs relative autonomi.
En problematiserings- og erfarings-analytisk undersøgelse*
23. Lichen Alex Yu
*Fabricating an S&OP Process
Circulating References and Matters of Concern*
24. Esben Alfort
*The Expression of a Need
Understanding search*
25. Trine Pallesen
*Assembling Markets for Wind Power
An Inquiry into the Making of Market Devices*
26. Anders Koed Madsen
*Web-Visions
Repurposing digital traces to organize social attention*
27. Lærke Højgaard Christiansen
BREWING ORGANIZATIONAL RESPONSES TO INSTITUTIONAL LOGICS
28. Tommy Kjær Lassen
*EGENTLIG SELVLEDELSE
En ledelsesfilosofisk afhandling om selvledelsens paradoksale dynamik og eksistentielle engagement*

29. Morten Rossing
Local Adaption and Meaning Creation in Performance Appraisal
30. Søren Obed Madsen
*Lederen som oversætter
Et oversættelsesteoretisk perspektiv på strategisk arbejde*
31. Thomas Høgenhaven
*Open Government Communities
Does Design Affect Participation?*
32. Kirstine Zinck Pedersen
*Failsafe Organizing?
A Pragmatic Stance on Patient Safety*
33. Anne Petersen
*Hverdagslogikker i psykiatrisk arbejde
En institutionsetnografisk undersøgelse af hverdagen i psykiatriske organisationer*
34. Dikke Maria Humle
Fortællinger om arbejde
35. Mark Holst-Mikkelsen
Strategieksekverering i praksis – barrierer og muligheder!
36. Malek Maalouf
*Sustaining lean
Strategies for dealing with organizational paradoxes*
37. Nicolaj Tofte Brenneche
*Systemic Innovation In The Making
The Social Productivity of Cartographic Crisis and Transitions in the Case of SEET*
38. Morten Gylling
*The Structure of Discourse
A Corpus-Based Cross-Linguistic Study*
39. Binzhang YANG
Urban Green Spaces for Quality Life - Case Study: the landscape architecture for people in Copenhagen
40. Michael Friis Pedersen
*Finance and Organization:
The Implications for Whole Farm Risk Management*
41. Even Fallan
Issues on supply and demand for environmental accounting information
42. Ather Nawaz
*Website user experience
A cross-cultural study of the relation between users' cognitive style, context of use, and information architecture of local websites*
43. Karin Beukel
The Determinants for Creating Valuable Inventions
44. Arjan Markus
*External Knowledge Sourcing and Firm Innovation
Essays on the Micro-Foundations of Firms' Search for Innovation*
- 2014**
1. Solon Moreira
Four Essays on Technology Licensing and Firm Innovation
2. Karin Strzeletz Ivertsen
*Partnership Drift in Innovation Processes
A study of the Think City electric car development*
3. Kathrine Hoffmann Pii
Responsibility Flows in Patient-centred Prevention
4. Jane Bjørn Vedel
*Managing Strategic Research
An empirical analysis of science-industry collaboration in a pharmaceutical company*
5. Martin Gylling
*Processuel strategi i organisationer
Monografi om dobbeltheden i tænkning af strategi, dels som vidensfelt i organisationsteori, dels som kunstnerisk tilgang til at skabe i erhvervsmæssig innovation*

6. Linne Marie Lauesen
Corporate Social Responsibility in the Water Sector: How Material Practices and their Symbolic and Physical Meanings Form a Colonising Logic
7. Maggie Qiuzhu Mei
LEARNING TO INNOVATE: The role of ambidexterity, standard, and decision process
8. Inger Høedt-Rasmussen
Developing Identity for Lawyers Towards Sustainable Lawyering
9. Sebastian Fux
Essays on Return Predictability and Term Structure Modelling
10. Thorbjørn N. M. Lund-Poulsen
Essays on Value Based Management
11. Oana Brindusa Albu
Transparency in Organizing: A Performative Approach
12. Lena Olaison
Entrepreneurship at the limits
13. Hanne Sørum
DRESSED FOR WEB SUCCESS? An Empirical Study of Website Quality in the Public Sector
14. Lasse Folke Henriksen
Knowing networks How experts shape transnational governance
15. Maria Halbinger
Entrepreneurial Individuals Empirical Investigations into Entrepreneurial Activities of Hackers and Makers
16. Robert Spliid
Kapitalfondenes metoder og kompetencer
17. Christiane Stelling
Public-private partnerships & the need, development and management of trusting A processual and embedded exploration
18. Marta Gasparin
Management of design as a translation process
19. Kåre Moberg
Assessing the Impact of Entrepreneurship Education From ABC to PhD
20. Alexander Cole
Distant neighbors Collective learning beyond the cluster
21. Martin Møller Boje Rasmussen
Is Competitiveness a Question of Being Alike? How the United Kingdom, Germany and Denmark Came to Compete through their Knowledge Regimes from 1993 to 2007
22. Anders Ravn Sørensen
Studies in central bank legitimacy, currency and national identity Four cases from Danish monetary history
23. Nina Bellak
Can Language be Managed in International Business? Insights into Language Choice from a Case Study of Danish and Austrian Multinational Corporations (MNCs)
24. Rikke Kristine Nielsen
Global Mindset as Managerial Meta-competence and Organizational Capability: Boundary-crossing Leadership Cooperation in the MNC The Case of 'Group Mindset' in Solar AIS.
25. Rasmus Koss Hartmann
User Innovation inside government Towards a critically performative foundation for inquiry

26. Kristian Gylling Olesen
Flertydig og emergerende ledelse i folkeskolen
Et aktør-netværksteoretisk ledelsesstudie af politiske evalueringsreformers betydning for ledelse i den danske folkeskole
27. Troels Riis Larsen
Kampen om Danmarks omdømme 1945-2010
Omdømmearbejde og omdømmepolitik
28. Klaus Majgaard
Jagten på autenticitet i offentlig styring
29. Ming Hua Li
Institutional Transition and Organizational Diversity: Differentiated internationalization strategies of emerging market state-owned enterprises
30. Sofie Blinkenberg Federspiel
IT, organisation og digitalisering: Institutionelt arbejde i den kommunale digitaliseringsproces
31. Elvi Weinreich
Hvilke offentlige ledere er der brug for når velfærdstænkningen flytter sig – er Diplomuddannelsens lederprofil svaret?
32. Ellen Mølgaard Korsager
Self-conception and image of context in the growth of the firm
– A Penrosian History of Fiberline Composites
33. Else Skjold
The Daily Selection
34. Marie Louise Conradsen
The Cancer Centre That Never Was
The Organisation of Danish Cancer Research 1949-1992
35. Virgilio Failla
Three Essays on the Dynamics of Entrepreneurs in the Labor Market
36. Nicky Nedergaard
Brand-Based Innovation
Relational Perspectives on Brand Logics and Design Innovation Strategies and Implementation
37. Mads Gjedsted Nielsen
Essays in Real Estate Finance
38. Kristin Martina Brandl
Process Perspectives on Service Offshoring
39. Mia Rosa Koss Hartmann
In the gray zone
With police in making space for creativity
40. Karen Ingerslev
Healthcare Innovation under The Microscope
Framing Boundaries of Wicked Problems
41. Tim Neerup Thomsen
Risk Management in large Danish public capital investment programmes
- 2015**
1. Jakob Ion Wille
Film som design
Design af levende billeder i film og tv-serier
2. Christiane Mossin
Interzones of Law and Metaphysics
Hierarchies, Logics and Foundations of Social Order seen through the Prism of EU Social Rights
3. Thomas Tøth
TRUSTWORTHINESS: ENABLING GLOBAL COLLABORATION
An Ethnographic Study of Trust, Distance, Control, Culture and Boundary Spanning within Offshore Outsourcing of IT Services
4. Steven Højlund
Evaluation Use in Evaluation Systems – The Case of the European Commission

5. Julia Kirch Kirkegaard
AMBIGUOUS WINDS OF CHANGE – OR FIGHTING AGAINST WINDMILLS IN CHINESE WIND POWER
A CONSTRUCTIVIST INQUIRY INTO CHINA'S PRAGMATICS OF GREEN MARKETISATION MAPPING
CONTROVERSIES OVER A POTENTIAL TURN TO QUALITY IN CHINESE WIND POWER
6. Michelle Carol Antero
A Multi-case Analysis of the Development of Enterprise Resource Planning Systems (ERP) Business Practices

Morten Friis-Olivarius
The Associative Nature of Creativity
7. Mathew Abraham
New Cooperativism: A study of emerging producer organisations in India
8. Stine Hedegaard
Sustainability-Focused Identity: Identity work performed to manage, negotiate and resolve barriers and tensions that arise in the process of constructing or ganizational identity in a sustainability context
9. Cecilie Glerup
Organizing Science in Society – the conduct and justification of resposable research
10. Allan Salling Pedersen
Implementering af ITIL® IT-governance - når best practice konflikter med kulturen Løsning af implementerings-problemer gennem anvendelse af kendte CSF i et aktionsforskningsforløb.
11. Nihat Misir
A Real Options Approach to Determining Power Prices
12. Mamdouh Medhat
MEASURING AND PRICING THE RISK OF CORPORATE FAILURES
13. Rina Hansen
Toward a Digital Strategy for Omnichannel Retailing
14. Eva Pallesen
In the rhythm of welfare creation
A relational processual investigation moving beyond the conceptual horizon of welfare management
15. Gouya Harirchi
In Search of Opportunities: Three Essays on Global Linkages for Innovation
16. Lotte Holck
Embedded Diversity: A critical ethnographic study of the structural tensions of organizing diversity
17. Jose Daniel Balarezo
Learning through Scenario Planning
18. Louise Pram Nielsen
Knowledge dissemination based on terminological ontologies. Using eye tracking to further user interface design.
19. Sofie Dam
PUBLIC-PRIVATE PARTNERSHIPS FOR INNOVATION AND SUSTAINABILITY TRANSFORMATION
An embedded, comparative case study of municipal waste management in England and Denmark
20. Ulrik Hartmyer Christiansen
Following the Content of Reported Risk Across the Organization
21. Guro Refsum Sanden
Language strategies in multinational corporations. A cross-sector study of financial service companies and manufacturing companies.
22. Linn Gevoll
Designing performance management for operational level
- A closer look on the role of design choices in framing coordination and motivation

23. Frederik Larsen
*Objects and Social Actions
– on Second-hand Valuation Practices*
24. Thorhildur Hansdottir Jetzek
*The Sustainable Value of Open
Government Data
Uncovering the Generative Mechanisms
of Open Data through a Mixed
Methods Approach*
25. Gustav Toppenberg
*Innovation-based M&A
– Technological-Integration
Challenges – The Case of
Digital-Technology Companies*
26. Mie Plotnikof
*Challenges of Collaborative
Governance
An Organizational Discourse Study
of Public Managers' Struggles
with Collaboration across the
Daycare Area*
27. Christian Garmann Johnsen
*Who Are the Post-Bureaucrats?
A Philosophical Examination of the
Creative Manager, the Authentic Leader
and the Entrepreneur*
28. Jacob Brogaard-Kay
*Constituting Performance Management
A field study of a pharmaceutical
company*
29. Rasmus Ploug Jenle
*Engineering Markets for Control:
Integrating Wind Power into the Danish
Electricity System*
30. Morten Lindholst
*Complex Business Negotiation:
Understanding Preparation and
Planning*
31. Morten Grynings
*TRUST AND TRANSPARENCY FROM AN
ALIGNMENT PERSPECTIVE*
32. Peter Andreas Norn
*Byregimer og styringsevne: Politisk
lederskab af store byudviklingsprojekter*
33. Milan Miric
*Essays on Competition, Innovation and
Firm Strategy in Digital Markets*
34. Sanne K. Hjordrup
*The Value of Talent Management
Rethinking practice, problems and
possibilities*
35. Johanna Sax
*Strategic Risk Management
– Analyzing Antecedents and
Contingencies for Value Creation*
36. Pernille Rydén
Strategic Cognition of Social Media
37. Mimmi Sjöklint
*The Measurable Me
- The Influence of Self-tracking on the
User Experience*
38. Juan Ignacio Staricco
*Towards a Fair Global Economic
Regime? A critical assessment of Fair
Trade through the examination of the
Argentinean wine industry*
39. Marie Henriette Madsen
*Emerging and temporary connections
in Quality work*
40. Yangfeng CAO
*Toward a Process Framework of
Business Model Innovation in the
Global Context
Entrepreneurship-Enabled Dynamic
Capability of Medium-Sized
Multinational Enterprises*
41. Carsten Scheibye
*Enactment of the Organizational Cost
Structure in Value Chain Configuration
A Contribution to Strategic Cost
Management*

2016

1. Signe Sofie Dyrby
Enterprise Social Media at Work
2. Dorte Boesby Dahl
*The making of the public parking attendant
Dirt, aesthetics and inclusion in public service work*
3. Verena Girschik
*Realizing Corporate Responsibility
Positioning and Framing in Nascent Institutional Change*
4. Anders Ørding Olsen
*IN SEARCH OF SOLUTIONS
Inertia, Knowledge Sources and Diversity in Collaborative Problem-solving*
5. Pernille Steen Pedersen
*Udkast til et nyt copingbegreb
En kvalifikation af ledelsesmuligheder for at forebygge sygefravær ved psykiske problemer.*
6. Kerli Kant Hvass
*Weaving a Path from Waste to Value:
Exploring fashion industry business models and the circular economy*
7. Kasper Lindskow
Exploring Digital News Publishing Business Models – a production network approach
8. Mikkel Mouritz Marfelt
*The chameleon workforce:
Assembling and negotiating the content of a workforce*
9. Marianne Bertelsen
*Aesthetic encounters
Rethinking autonomy, space & time in today's world of art*
10. Louise Hauberg Wilhelmsen
EU PERSPECTIVES ON INTERNATIONAL COMMERCIAL ARBITRATION
11. Abid Hussain
On the Design, Development and Use of the Social Data Analytics Tool (SODATO): Design Propositions, Patterns, and Principles for Big Social Data Analytics
12. Mark Bruun
Essays on Earnings Predictability
13. Tor Bøe-Lillegraven
BUSINESS PARADOXES, BLACK BOXES, AND BIG DATA: BEYOND ORGANIZATIONAL AMBIDEXTERITY
14. Hadis Khonsary-Atighi
ECONOMIC DETERMINANTS OF DOMESTIC INVESTMENT IN AN OIL-BASED ECONOMY: THE CASE OF IRAN (1965-2010)
15. Maj Lervad Grasten
*Rule of Law or Rule by Lawyers?
On the Politics of Translation in Global Governance*
16. Lene Granzau Juel-Jacobsen
SUPERMARKEDETS MODUS OPERANDI – en hverdagssociologisk undersøgelse af forholdet mellem rum og handlen og understøtte relationsopbygning?
17. Christine Thalsgård Henriques
In search of entrepreneurial learning – Towards a relational perspective on incubating practices?
18. Patrick Bennett
Essays in Education, Crime, and Job Displacement
19. Søren Korsgaard
Payments and Central Bank Policy
20. Marie Kruse Skibsted
Empirical Essays in Economics of Education and Labor
21. Elizabeth Benedict Christensen
*The Constantly Contingent Sense of Belonging of the 1.5 Generation
Undocumented Youth
An Everyday Perspective*

22. Lasse J. Jessen
Essays on Discounting Behavior and Gambling Behavior
23. Kalle Johannes Rose
*Når stiftertiljen dør...
Et retsøkonomisk bidrag til 200 års
juridisk konflikt om ejendomsretten*
24. Andreas Søeborg Kirkedal
*Danish Stød and Automatic Speech
Recognition*
25. Ida Lunde Jørgensen
*Institutions and Legitimations in
Finance for the Arts*
26. Olga Rykov Ibsen
*An empirical cross-linguistic study of
directives: A semiotic approach to the
sentence forms chosen by British,
Danish and Russian speakers in native
and ELF contexts*
27. Desi Volker
Understanding Interest Rate Volatility
28. Angeli Elizabeth Weller
*Practice at the Boundaries of Business
Ethics & Corporate Social Responsibility*
29. Ida Danneskiold-Samsøe
*Levende læring i kunstneriske
organisationer
En undersøgelse af læringsprocesser
mellem projekt og organisation på
Aarhus Teater*
30. Leif Christensen
*Quality of information – The role of
internal controls and materiality*
31. Olga Zarzecka
Tie Content in Professional Networks
32. Henrik Mahncke
*De store gaver
- Filantropiens gensidighedsrelationer i
teori og praksis*
33. Carsten Lund Pedersen
*Using the Collective Wisdom of
Frontline Employees in Strategic Issue
Management*
34. Yun Liu
Essays on Market Design
35. Denitsa Hazarbassanova Blagoeva
The Internationalisation of Service Firms
36. Manya Jaura Lind
*Capability development in an off-
shoring context: How, why and by
whom*
37. Luis R. Boscán F.
*Essays on the Design of Contracts and
Markets for Power System Flexibility*
38. Andreas Philipp Distel
*Capabilities for Strategic Adaptation:
Micro-Foundations, Organizational
Conditions, and Performance
Implications*
39. Lavinia Bleoca
*The Usefulness of Innovation and
Intellectual Capital in Business
Performance: The Financial Effects of
Knowledge Management vs. Disclosure*
40. Henrik Jensen
*Economic Organization and Imperfect
Managerial Knowledge: A Study of the
Role of Managerial Meta-Knowledge
in the Management of Distributed
Knowledge*
41. Stine Mosekjær
*The Understanding of English Emotion
Words by Chinese and Japanese
Speakers of English as a Lingua Franca
An Empirical Study*
42. Hallur Tor Sigurdarson
*The Ministry of Desire - Anxiety and
entrepreneurship in a bureaucracy*
43. Kätlin Pulk
*Making Time While Being in Time
A study of the temporality of
organizational processes*
44. Valeria Giacomini
*Contextualizing the cluster Palm oil in
Southeast Asia in global perspective
(1880s–1970s)*

45. Jeanette Willert
Managers' use of multiple Management Control Systems: The role and interplay of management control systems and company performance
46. Mads Vestergaard Jensen
Financial Frictions: Implications for Early Option Exercise and Realized Volatility
47. Mikael Reimer Jensen
Interbank Markets and Frictions
48. Benjamin Faigen
Essays on Employee Ownership
49. Adela Michea
Enacting Business Models An Ethnographic Study of an Emerging Business Model Innovation within the Frame of a Manufacturing Company.
50. Iben Sandal Stjerne
Transcending organization in temporary systems Aesthetics' organizing work and employment in Creative Industries
51. Simon Krogh
Anticipating Organizational Change
52. Sarah Netter
Exploring the Sharing Economy
53. Lene Tolstrup Christensen
State-owned enterprises as institutional market actors in the marketization of public service provision: A comparative case study of Danish and Swedish passenger rail 1990–2015
54. Kyoung(Kay) Sun Park
Three Essays on Financial Economics
- 2017**
1. Mari Bjerck
Apparel at work. Work uniforms and women in male-dominated manual occupations.
2. Christoph H. Flöthmann
Who Manages Our Supply Chains? Backgrounds, Competencies and Contributions of Human Resources in Supply Chain Management
3. Aleksandra Anna Rzeźnik
Essays in Empirical Asset Pricing
4. Claes Bäckman
Essays on Housing Markets
5. Kirsti Reitan Andersen
Stabilizing Sustainability in the Textile and Fashion Industry
6. Kira Hoffmann
Cost Behavior: An Empirical Analysis of Determinants and Consequences of Asymmetries
7. Tobin Hanspal
Essays in Household Finance
8. Nina Lange
Correlation in Energy Markets
9. Anjum Fayyaz
Donor Interventions and SME Networking in Industrial Clusters in Punjab Province, Pakistan
10. Magnus Paulsen Hansen
Trying the unemployed. Justification and critique, emancipation and coercion towards the 'active society'. A study of contemporary reforms in France and Denmark
11. Sameer Azizi
Corporate Social Responsibility in Afghanistan – a critical case study of the mobile telecommunications industry

- | | |
|---|--|
| <p>12. Malene Myhre
<i>The internationalization of small and medium-sized enterprises: A qualitative study</i></p> | <p>23. Simone Stæhr
<i>Financial Analysts' Forecasts Behavioral Aspects and the Impact of Personal Characteristics</i></p> |
| <p>13. Thomas Presskorn-Thygesen
<i>The Significance of Normativity – Studies in Post-Kantian Philosophy and Social Theory</i></p> | <p>24. Mikkel Godt Gregersen
<i>Management Control, Intrinsic Motivation and Creativity – How Can They Coexist</i></p> |
| <p>14. Federico Clementi
<i>Essays on multinational production and international trade</i></p> | <p>25. Kristjan Johannes Suse Jespersen
<i>Advancing the Payments for Ecosystem Service Discourse Through Institutional Theory</i></p> |
| <p>15. Lara Anne Hale
<i>Experimental Standards in Sustainability Transitions: Insights from the Building Sector</i></p> | <p>26. Kristian Bondo Hansen
<i>Crowds and Speculation: A study of crowd phenomena in the U.S. financial markets 1890 to 1940</i></p> |
| <p>16. Richard Pucci
<i>Accounting for Financial Instruments in an Uncertain World Controversies in IFRS in the Aftermath of the 2008 Financial Crisis</i></p> | <p>27. Lars Balslev
<i>Actors and practices – An institutional study on management accounting change in Air Greenland</i></p> |
| <p>17. Sarah Maria Denta
<i>Kommunale offentlige private partnerskaber Regulering I skyggen af Farumsagen</i></p> | <p>28. Sven Klingler
<i>Essays on Asset Pricing with Financial Frictions</i></p> |
| <p>18. Christian Östlund
<i>Design for e-training</i></p> | <p>29. Klement Ahrensbach Rasmussen
<i>Business Model Innovation The Role of Organizational Design</i></p> |
| <p>19. Amalie Martinus Hauge
<i>Organizing Valuations – a pragmatic inquiry</i></p> | <p>30. Giulio Zichella
<i>Entrepreneurial Cognition. Three essays on entrepreneurial behavior and cognition under risk and uncertainty</i></p> |
| <p>20. Tim Holst Celik
<i>Tension-filled Governance? Exploring the Emergence, Consolidation and Reconfiguration of Legitimatory and Fiscal State-crafting</i></p> | <p>31. Richard Ledborg Hansen
<i>En forkærlighed til det eksisterende – mellemlederens oplevelse af forandringsmodstand i organisatoriske forandringer</i></p> |
| <p>21. Christian Bason
<i>Leading Public Design: How managers engage with design to transform public governance</i></p> | <p>32. Vilhelm Stefan Holsting
<i>Militært chefvirke: Kritik og retfærdiggørelse mellem politik og profession</i></p> |
| <p>22. Davide Tomio
<i>Essays on Arbitrage and Market Liquidity</i></p> | |

33. Thomas Jensen **2018**
Shipping Information Pipeline: An information infrastructure to improve international containerized shipping
34. Dzmitry Bartalevich
Do economic theories inform policy? Analysis of the influence of the Chicago School on European Union competition policy
35. Kristian Roed Nielsen
Crowdfunding for Sustainability: A study on the potential of reward-based crowdfunding in supporting sustainable entrepreneurship
36. Emil Husted
There is always an alternative: A study of control and commitment in political organization
37. Anders Ludvig Sevelsted
Interpreting Bonds and Boundaries of Obligation. A genealogy of the emergence and development of Protestant voluntary social work in Denmark as shown through the cases of the Copenhagen Home Mission and the Blue Cross (1850 – 1950)
38. Niklas Kohl
Essays on Stock Issuance
39. Maya Christiane Flensburg Jensen
BOUNDARIES OF PROFESSIONALIZATION AT WORK An ethnography-inspired study of care workers' dilemmas at the margin
40. Andreas Kamstrup
Crowdsourcing and the Architectural Competition as Organisational Technologies
41. Louise Lyngfeldt Gorm Hansen
Triggering Earthquakes in Science, Politics and Chinese Hydropower - A Controversy Study
1. Vishv Priya Kohli
Combatting Falsification and Counterfeiting of Medicinal Products in the European Union – A Legal Analysis
2. Helle Haurum
Customer Engagement Behavior in the context of Continuous Service Relationships
3. Nis Grünberg
The Party-state order: Essays on China's political organization and political economic institutions
4. Jesper Christensen
A Behavioral Theory of Human Capital Integration
5. Poula Marie Helth
Learning in practice
6. Rasmus Vendler Toft-Kehler
Entrepreneurship as a career? An investigation of the relationship between entrepreneurial experience and entrepreneurial outcome
7. Szymon Furtak
Sensing the Future: Designing sensor-based predictive information systems for forecasting spare part demand for diesel engines
8. Mette Brehm Johansen
Organizing patient involvement. An ethnographic study
9. Iwona Sulinska
Complexities of Social Capital in Boards of Directors
10. Cecilie Fanø Petersen
Award of public contracts as a means to conferring State aid: A legal analysis of the interface between public procurement law and State aid law
11. Ahmad Ahmad Barirani
Three Experimental Studies on Entrepreneurship

12. Carsten Allerslev Olsen
Financial Reporting Enforcement: Impact and Consequences
13. Irene Christensen
New product fumbles – Organizing for the Ramp-up process
14. Jacob Taarup-Esbensen
Managing communities – Mining MNEs' community risk management practices
15. Lester Allan Lasrado
Set-Theoretic approach to maturity models
16. Mia B. Münster
Intention vs. Perception of Designed Atmospheres in Fashion Stores
17. Anne Sluhan
Non-Financial Dimensions of Family Firm Ownership: How Socioemotional Wealth and Familiness Influence Internationalization
18. Henrik Yde Andersen
Essays on Debt and Pensions
19. Fabian Heinrich Müller
Valuation Reversed – When Valuers are Valuated. An Analysis of the Perception of and Reaction to Reviewers in Fine-Dining
20. Martin Jarmatz
Organizing for Pricing
21. Niels Joachim Christfort Gormsen
Essays on Empirical Asset Pricing
22. Diego Zunino
Socio-Cognitive Perspectives in Business Venturing
23. Benjamin Asmussen
Networks and Faces between Copenhagen and Canton, 1730-1840
24. Dalia Bagdziunaite
Brains at Brand Touchpoints A Consumer Neuroscience Study of Information Processing of Brand Advertisements and the Store Environment in Compulsive Buying
25. Erol Kazan
Towards a Disruptive Digital Platform Model
26. Andreas Bang Nielsen
Essays on Foreign Exchange and Credit Risk
27. Anne Krebs
Accountable, Operable Knowledge Toward Value Representations of Individual Knowledge in Accounting
28. Matilde Fogh Kirkegaard
A firm- and demand-side perspective on behavioral strategy for value creation: Insights from the hearing aid industry
29. Agnieszka Nowinska
SHIPS AND RELATION-SHIPS Tie formation in the sector of shipping intermediaries in shipping
30. Stine Evald Bentsen
The Comprehension of English Texts by Native Speakers of English and Japanese, Chinese and Russian Speakers of English as a Lingua Franca. An Empirical Study.
31. Stine Louise Daetz
Essays on Financial Frictions in Lending Markets
32. Christian Skov Jensen
Essays on Asset Pricing
33. Anders Kryger
Aligning future employee action and corporate strategy in a resource-scarce environment

34. Maitane Elorriaga-Rubio
The behavioral foundations of strategic decision-making: A contextual perspective
35. Roddy Walker
Leadership Development as Organisational Rehabilitation: Shaping Middle-Managers as Double Agents
36. Jinsun Bae
Producing Garments for Global Markets Corporate social responsibility (CSR) in Myanmar's export garment industry 2011–2015
37. Queralt Prat-i-Pubill
Axiological knowledge in a knowledge driven world. Considerations for organizations.
38. Pia Mølgaard
Essays on Corporate Loans and Credit Risk
39. Marzia Aricò
Service Design as a Transformative Force: Introduction and Adoption in an Organizational Context
40. Christian Dyrland Wåhlin-Jacobsen
Constructing change initiatives in workplace voice activities Studies from a social interaction perspective
41. Peter Kalum Schou
Institutional Logics in Entrepreneurial Ventures: How Competing Logics arise and shape organizational processes and outcomes during scale-up
42. Per Henriksen
Enterprise Risk Management Rationaler og paradokser i en moderne ledelsesteknologi
43. Maximilian Schellmann
The Politics of Organizing Refugee Camps
44. Jacob Halvas Bjerre
Excluding the Jews: The Aryanization of Danish-German Trade and German Anti-Jewish Policy in Denmark 1937-1943
45. Ida Schrøder
Hybridising accounting and caring: A symmetrical study of how costs and needs are connected in Danish child protection work
46. Katrine Kunst
Electronic Word of Behavior: Transforming digital traces of consumer behaviors into communicative content in product design
47. Viktor Avlonitis
Essays on the role of modularity in management: Towards a unified perspective of modular and integral design
48. Anne Sofie Fischer
Negotiating Spaces of Everyday Politics: -An ethnographic study of organizing for social transformation for women in urban poverty, Delhi, India

2019

1. Shihan Du
*ESSAYS IN EMPIRICAL STUDIES
BASED ON ADMINISTRATIVE
LABOUR MARKET DATA*
2. Mart Laatsit
*Policy learning in innovation
policy: A comparative analysis of
European Union member states*
3. Peter J. Wynne
*Proactively Building Capabilities for
the Post-Acquisition Integration
of Information Systems*
4. Kalina S. Staykova
*Generative Mechanisms for Digital
Platform Ecosystem Evolution*
5. Ieva Linkeviciute
*Essays on the Demand-Side
Management in Electricity Markets*
6. Jonatan Echebarria Fernández
*Jurisdiction and Arbitration
Agreements in Contracts for the
Carriage of Goods by Sea –
Limitations on Party Autonomy*
7. Louise Thorn Bøttkjær
*Votes for sale. Essays on
clientelism in new democracies.*
8. Ditte Vilstrup Holm
*The Poetics of Participation:
the organizing of participation in
contemporary art*
9. Philip Rosenbaum
*Essays in Labor Markets –
Gender, Fertility and Education*
10. Mia Olsen
*Mobile Betalinger - Succesfaktorer
og Adfærdsmæssige Konsekvenser*
11. Adrián Luis Mérida Gutiérrez
*Entrepreneurial Careers:
Determinants, Trajectories, and
Outcomes*
12. Frederik Regli
Essays on Crude Oil Tanker Markets
13. Cancan Wang
*Becoming Adaptive through Social
Media: Transforming Governance and
Organizational Form in Collaborative
E-government*
14. Lena Lindbjerg Sperling
*Economic and Cultural Development:
Empirical Studies of Micro-level Data*
15. Xia Zhang
*Obligation, face and facework:
An empirical study of the communi-
cative act of cancellation of an
obligation by Chinese, Danish and
British business professionals in both
L1 and ELF contexts*
16. Stefan Kirkegaard Sløk-Madsen
*Entrepreneurial Judgment and
Commercialization*
17. Erin Leitheiser
*The Comparative Dynamics of Private
Governance
The case of the Bangladesh Ready-
Made Garment Industry*
18. Lone Christensen
*STRATEGIIMPLEMENTERING:
STYRINGSBESTRÆBELSER, IDENTITET
OG AFFEKT*
19. Thomas Kjær Poulsen
*Essays on Asset Pricing with Financial
Frictions*
20. Maria Lundberg
*Trust and self-trust in leadership iden-
tity constructions: A qualitative explo-
ration of narrative ecology in the dis-
cursive aftermath of heroic discourse*

21. Tina Joanes
*Sufficiency for sustainability
Determinants and strategies for reducing
clothing consumption*
22. Benjamin Johannes Flesch
*Social Set Visualizer (SoSeVi): Design,
Development and Evaluation of a Visual
Analytics Tool for Computational Set
Analysis of Big Social Data*
23. Henriette Sophia Groskopff
Tvede Schleimann
*Creating innovation through collaboration
– Partnering in the maritime sector*
24. Kristian Steensen Nielsen
*The Role of Self-Regulation in
Environmental Behavior Change*
25. Lydia L. Jørgensen
Moving Organizational Atmospheres
26. Theodor Lucian Vladasel
*Embracing Heterogeneity: Essays in
Entrepreneurship and Human Capital*
27. Seidi Suurmets
*Contextual Effects in Consumer Research:
An Investigation of Consumer Information
Processing and Behavior via the Applicati
on of Eye-tracking Methodology*
28. Marie Sundby Palle Nickelsen
*Reformer mellem integritet og innovation:
Reform af reformens form i den danske
centraladministration fra 1920 til 2019*
29. Vibeke Kristine Scheller
*The temporal organizing of same-day
discharge: A tempography of a Cardiac
Day Unit*
30. Qian Sun
*Adopting Artificial Intelligence in
Healthcare in the Digital Age: Perceived
Challenges, Frame Incongruence, and
Social Power*
31. Dorthe Thorning Mejlhede
*Artful change agency and organizing for
innovation – the case of a Nordic fintech
cooperative*
32. Benjamin Christoffersen
*Corporate Default Models:
Empirical Evidence and Methodical
Contributions*
33. Filipe Antonio Bonito Vieira
Essays on Pensions and Fiscal Sustainability
34. Morten Nicklas Bigler Jensen
*Earnings Management in Private Firms:
An Empirical Analysis of Determinants
and Consequences of Earnings
Management in Private Firms*
- 2020**
1. Christian Hendriksen
*Inside the Blue Box: Explaining industry
influence in the International Maritime
Organization*
2. Vasileios Kosmas
*Environmental and social issues in global
supply chains:
Emission reduction in the maritime
transport industry and maritime search and
rescue operational response to migration*
3. Thorben Peter Simonsen
*The spatial organization of psychiatric
practice: A situated inquiry into 'healing
architecture'*
4. Signe Bruskin
*The infinite storm: An ethnographic study
of organizational change in a bank*
5. Rasmus Corlin Christensen
*Politics and Professionals: Transnational
Struggles to Change International Taxation*
6. Robert Lorenz Törmer
*The Architectural Enablement of a Digital
Platform Strategy*

7. Anna Kirkebæk Johansson Gosovic
Ethics as Practice: An ethnographic study of business ethics in a multinational biopharmaceutical company
8. Frank Meier
Making up leaders in leadership development
9. Kai Basner
Servitization at work: On proliferation and containment
10. Anestis Keremis
Anti-corruption in action: How is anti-corruption practiced in multinational companies?
11. Marie Larsen Ryberg
Governing Interdisciplinarity: Stakes and translations of interdisciplinarity in Danish high school education.
12. Jannick Friis Christensen
Queering organisation(s): Norm-critical orientations to organising and researching diversity
13. Thorsteinn Sigurdur Sveinsson
Essays on Macroeconomic Implications of Demographic Change
14. Catherine Casler
Reconstruction in strategy and organization: For a pragmatic stance
15. Luisa Murphy
Revisiting the standard organization of multi-stakeholder initiatives (MSIs): The case of a meta-MSI in Southeast Asia
16. Friedrich Bergmann
Essays on International Trade
17. Nicholas Haagensen
European Legal Networks in Crisis: The Legal Construction of Economic Policy
18. Charlotte Biil
Samskabelse med en sommerfugle-model: Hybrid ret i forbindelse med et partnerskabsprojekt mellem 100 selvejende daginstitutioner, deres paraplyorganisation, tre kommuner og CBS
19. Andreas Dimmelmeier
The Role of Economic Ideas in Sustainable Finance: From Paradigms to Policy
20. Maibrith Kempka Jensen
Ledelse og autoritet i interaktion - En interaktionsbaseret undersøgelse af autoritet i ledelse i praksis
21. Thomas Burø
LAND OF LIGHT: Assembling the Ecology of Culture in Odsherred 2000-2018
22. Prins Marcus Valiant Lantz
Timely Emotion: The Rhetorical Framing of Strategic Decision Making
23. Thorbjørn Vittenhof Fejerskov
Fra værdi til invitationer - offentlig værdiskabelse gennem affekt, potentialitet og begivenhed
24. Lea Acre Foverskov
Demographic Change and Employment: Path dependencies and institutional logics in the European Commission
25. Anirudh Agrawal
A Doctoral Dissertation
26. Julie Marx
Households in the housing market
27. Hadar Gafni
Alternative Digital Methods of Providing Entrepreneurial Finance

28. Mathilde Hjerrild Carlsen
Ledelse af engagementer: En undersøgelse af samarbejde mellem folkeskoler og virksomheder i Danmark
29. Suen Wang
Essays on the Gendered Origins and Implications of Social Policies in the Developing World
30. Stine Hald Larsen
The Story of the Relative: A Systems-Theoretical Analysis of the Role of the Relative in Danish Eldercare Policy from 1930 to 2020
31. Christian Casper Hofma
Immersive technologies and organizational routines: When head-mounted displays meet organizational routines
32. Jonathan Feddersen
The temporal emergence of social relations: An event-based perspective of organising
33. Nageswaran Vaidyanathan
ENRICHING RETAIL CUSTOMER EXPERIENCE USING AUGMENTED REALITY
05. Fei Liu
Emergent Technology Use in Consumer Decision Journeys: A Process-as-Propensity Approach
06. Jakob Rømer Barfod
Ledelse i militære højrisikoteams
07. Elham Shafiei Gol
Creative Crowdswork Arrangements
08. Árni Jóhan Petersen
Collective Imaginary as (Residual) Fantasy: A Case Study of the Faroese Oil Bonanza
09. Søren Bering
"Manufacturing, Forward Integration and Governance Strategy"
10. Lars Oehler
Technological Change and the Decomposition of Innovation: Choices and Consequences for Latecomer Firm Upgrading: The Case of China's Wind Energy Sector
11. Lise Dahl Arvedsen
Leadership in interaction in a virtual context: A study of the role of leadership processes in a complex context, and how such processes are accomplished in practice

2021

1. Vanya Rusinova
The Determinants of Firms' Engagement in Corporate Social Responsibility: Evidence from Natural Experiments
2. Lívia Lopes Barakat
Knowledge management mechanisms at MNCs: The enhancing effect of absorptive capacity and its effects on performance and innovation
3. Søren Bundgaard Brøgger
Essays on Modern Derivatives Markets
4. Martin Friis Nielsen
Consuming Memory: Towards a conceptualization of social media platforms as organizational technologies of consumption
12. Jacob Emil Jeppesen
Essays on Knowledge networks, scientific impact and new knowledge adoption
13. Kasper Ingeman Beck
Essays on Chinese State-Owned Enterprises: Reform, Corporate Governance and Subnational Diversity
14. Sönnich Dahl Sönnichsen
Exploring the interface between public demand and private supply for implementation of circular economy principles
15. Benjamin Knox
Essays on Financial Markets and Monetary Policy

16. Anita Eskesen
Essays on Utility Regulation: Evaluating Negotiation-Based Approaches in the Context of Danish Utility Regulation
17. Agnes Guenther
Essays on Firm Strategy and Human Capital
18. Sophie Marie Cappelen
Walking on Eggshells: The balancing act of temporal work in a setting of culinary change
19. Manar Saleh Alnamlah
About Gender Gaps in Entrepreneurial Finance
20. Kirsten Tangaa Nielsen
Essays on the Value of CEOs and Directors
21. Renée Ridgway
Re:search - the Personalised Subject vs. the Anonymous User
22. Codrina Ana Maria Lauth
IMPACT Industrial Hackathons: Findings from a longitudinal case study on short-term vs long-term IMPACT implementations from industrial hackathons within Grundfos
23. Wolf-Hendrik Uhlbach
Scientist Mobility: Essays on knowledge production and innovation
24. Tomaz Sedej
Blockchain technology and inter-organizational relationships
25. Lasse Bundgaard
Public Private Innovation Partnerships: Creating Public Value & Scaling Up Sustainable City Solutions
26. Dimitra Makri Andersen
Walking through Temporal Walls: Rethinking NGO Organizing for Sustainability through a Temporal Lens on NGO-Business Partnerships
27. Louise Fjord Kjærsgaard
Allocation of the Right to Tax Income from Digital Products and Services: A legal analysis of international tax treaty law
28. Sara Dahlman
Marginal alternativity: Organizing for sustainable investing
29. Henrik Gundelach
Performance determinants: An Investigation of the Relationship between Resources, Experience and Performance in Challenging Business Environments
30. Tom Wraight
Confronting the Developmental State: American Trade Policy in the Neoliberal Era
31. Mathias Fjællegaard Jensen
Essays on Gender and Skills in the Labour Market
32. Daniel Lundgaard
Using Social Media to Discuss Global Challenges: Case Studies of the Climate Change Debate on Twitter
33. Jonas Sveistrup Søgaard
Designs for Accounting Information Systems using Distributed Ledger Technology
34. Sarosh Asad
CEO narcissism and board composition: Implications for firm strategy and performance
35. Johann Ole Willers
Experts and Markets in Cybersecurity On Definitional Power and the Organization of Cyber Risks
36. Alexander Kronies
Opportunities and Risks in Alternative Investments

37. Niels Fuglsang
The Politics of Economic Models: An inquiry into the possibilities and limits concerning the rise of macroeconomic forecasting models and what this means for policymaking

38. David Howoldt
Policy Instruments and Policy Mixes for Innovation: Analysing Their Relation to Grand Challenges, Entrepreneurship and Innovation Capability with Natural Language Processing and Latent Variable Methods

2022

01. Ditte Thøgersen
Managing Public Innovation on the Frontline

02. Rasmus Jørgensen
Essays on Empirical Asset Pricing and Private Equity

03. Nicola Giommetti
Essays on Private Equity

04. Laila Starr
When Is Health Innovation Worth It? Essays On New Approaches To Value Creation In Health

05. Maria Krysfeldt Rasmussen
Den transformative ledelsesbyrde – etnografisk studie af en religionsinspireret ledelsesfilosofi i en dansk modevirksomhed

06. Rikke Sejer Nielsen
Mortgage Decisions of Households: Consequences for Consumption and Savings

07. Myriam Noémy Marending
Essays on development challenges of low income countries: Evidence from conflict, pest and credit

08. Selorm Agbleze
A BEHAVIORAL THEORY OF FIRM FORMALIZATION

09. Rasmus Arler Bogetoft
Rettighedshavers faktisk lidte tab i immaterialretssager: Studier af dansk ret med støtte i økonomisk teori og metode

10. Franz Maximilian Buchmann
Driving the Green Transition of the Maritime Industry through Clean Technology Adoption and Environmental Policies

11. Ivan Olav Vulchanov
The role of English as an organisational language in international workplaces

12. Anne Agerbak Bilde
TRANSFORMATIONER AF SKOLELEDELSE - en systemteoretisk analyse af hvordan betingelser for skoleledelse forandres med læring som genstand i perioden 1958-2020

13. JUAN JOSE PRICE ELTON
EFFICIENCY AND PRODUCTIVITY ANALYSIS: TWO EMPIRICAL APPLICATIONS AND A METHODOLOGICAL CONTRIBUTION

14. Catarina Pessanha Gomes
The Art of Occupying: Romanticism as Political Culture in French Prefigurative politics

15. Mark Ørberg
Fondsretten og den levende vedtægt

16. Majbritt Greve
Maersk's Role in Economic Development: A Study of Shipping and Logistics Foreign Direct Investment in Global Trade

17. Sille Julie J. Abildgaard
Doing-Being Creative: Empirical Studies of Interaction in Design Work

18. Jette Sandager
Glitter, Glamour, and the Future of (More) Girls in STEM: Gendered Formations of STEM Aspirations

19. Casper Hein Winther
Inside the innovation lab - How paradoxical tensions persist in ambidextrous organizations over time

20. Nikola Kostić
Collaborative governance of inter-organizational relationships: The effects of management controls, blockchain technology, and industry standards
21. Salla Naomi Stausholm
Maximum capital, minimum tax: Enablers and facilitators of corporate tax minimization
22. Robin Porsfelt
Seeing through Signs: On Economic Imagination and Semiotic Speculation
23. Michael Herburger
Supply chain resilience – a concept for coping with cyber risks
24. Katharina Christiane Nielsen Jeschke
Balancing safety in everyday work - A case study of construction managers' dynamic safety practices
25. Jakob Ahm Sørensen
Financial Markets with Frictions and Belief Distortions
26. Jakob Laage-Thomsen
Nudging Leviathan, Protecting Demos - A Comparative Sociology of Public Administration and Expertise in the Nordics
27. Kathrine Søs Jacobsen Cesko
Collaboration between Economic Operators in the Competition for Public Contracts: A Legal and Economic Analysis of Grey Zones between EU Public Procurement Law and EU Competition Law
28. Mette Nelund
Den nye jord – Et feltstudie af et bæredygtigt virke på Farendløse Mosteri
29. Benjamin Cedric Larsen
Governing Artificial Intelligence – Lessons from the United States and China
30. Anders Brøndum Klein
Kollektiv meningsdannelse iblandt heterogene aktører i eksperimentelle samskabelsesprocesser
31. Stefano Tripodi
Essays on Development Economicis
32. Katrine Maria Lumbye
Internationalization of European Electricity Multinationals in Times of Transition
33. Xiaochun Guo
Dynamic Roles of Digital Currency – An Exploration from Interactive Processes: Difference, Time, and Perspective
34. Louise Lindbjerg
Three Essays on Firm Innovation
35. Marcela Galvis Restrepo
Feature reduction for classification with mixed data: an algorithmic approach
36. Hanna Nyborg Storm
Cultural institutions and attractiveness – How cultural institutions contribute to the development of regions and local communities
37. Anna-Bertha Heeris Christensen
Conflicts and Challenges in Practices of Commercializing Humans – An Ethnographic Study of Influencer Marketing Work
38. Casper Berg Lavmand Larsen
A Worker-Centered Inquiry into the Contingencies and Consequences of Worker Representation
39. Niels le Duc
The Resource Commitment of Multinational Enterprise R&D Activities
40. Esben Langager Olsen
Change management tools and change managers – Examining the simulacra of change
41. Anne Sophie Lassen
Gender in the Labor Market

42. Alison E. Holm
Corrective corporate responses to accusations of misconduct on societal issues
43. Chenyan Lyu
Carbon Pricing, Renewable Energy, and Clean Growth – A Market Perspective
44. Alina Grecu
UNPACKING MULTI-LEVEL OFFSHORING CONSEQUENCES: Hiring Wages, Onshore Performance, and Public Sentiment
45. Alexandra Lüth
Offshore Energy Hubs as an Emerging Concept – Sector Integration at Sea

2023

01. Cheryl Basil Sequeira
Port Business Development – Digitalisation of Port Authority and Hybrid Governance Model
02. Mette Suder Franck
Empirical Essays on Technology Supported Learning – Studies of Danish Higher Education
03. Søren Lund Frandsen
States and Experts – Assembling Expertise for Climate Change and Pandemics
04. Guowei Dong
Innovation and Internationalization – Evidence from Chinese Manufacturing Enterprises
05. Eileen Murphy
In Service to Security – Constructing the Authority to Manage European Border Data Infrastructures
06. Bontu Lucie Guschke
THE PERSISTENCE OF SEXISM AND RACISM AT UNIVERSITIES – Exploring the imperceptibility and unspeakability of workplace harassment and discrimination in academia
07. Christoph Viebig
Learning Entrepreneurship – How capabilities shape learning from experience, reflection, and action
08. Kasper Regenburt
Financial Risks of Private Firms
09. Kathrine Møller Solgaard
Who to hire? – A situated study of employee selection as routine, practice, and process
10. Jack Kværnø-Jones
Intersections between FinTech Imaginaries and Traditional Banking – A study of disciplinary, implementary, and parasitic work in the Danish financial sector
11. Stine Quorning
Managing Climate Change Like a Central Banker – The Political Economy of Greening the Monetary Technocracy
12. Amanda Bille
No business without politics – Investigating the political nature of supply chain management
13. Theis Ingerslev Jensen
Essays on Empirical Asset Pricing

TITLER I ATV PH.D.-SERIEN

1992

1. Niels Kornum
Servicemønstre – organisation, økonomi og planlægningsmetode

1995

2. Verner Worm
*Nordiske virksomheder i Kina
Kulturspecifikke interaktionsrelationer ved nordiske virksomhedsetableringer i Kina*

1999

3. Mogens Bjerre
*Key Account Management of Complex Strategic Relationships
An Empirical Study of the Fast Moving Consumer Goods Industry*

2000

4. Lotte Darsø
*Innovation in the Making
Interaction Research with heterogeneous Groups of Knowledge Workers creating new Knowledge and new Leads*

2001

5. Peter Hobolt Jensen
*Managing Strategic Design Identities
The case of the Lego Developer Network*

2002

6. Peter Lohmann
The Deleuzian Other of Organizational Change – Moving Perspectives of the Human
7. Anne Marie Jess Hansen
To lead from a distance: The dynamic interplay between strategy and strategizing – A case study of the strategic management process

2003

8. Lotte Henriksen
*Videndeling
– om organisatoriske og ledelsesmæssige udfordringer ved videndeling i praksis*
9. Niels Christian Nickelsen
Arrangements of Knowing: Coordinating Procedures Tools and Bodies in Industrial Production – a case study of the collective making of new products

2005

10. Carsten Ørts Hansen
Konstruktion af ledelsesteknologier og effektivitet

TITLER I DBA PH.D.-SERIEN

2007

1. Peter Kastrup-Misir
Endeavoring to Understand Market Orientation – and the concomitant co-mutation of the researched, the researcher, the research itself and the truth

2009

1. Torkild Leo Thellefsen
*Fundamental Signs and Significance effects
A Semeiotic outline of Fundamental Signs, Significance-effects, Knowledge Profiling and their use in Knowledge Organization and Branding*
2. Daniel Ronzani
When Bits Learn to Walk Don't Make Them Trip. Technological Innovation and the Role of Regulation by Law in Information Systems Research: the Case of Radio Frequency Identification (RFID)

2010

1. Alexander Carnera
*Magten over livet og livet som magt
Studier i den biopolitiske ambivalens*