Essays in Factor Investing

Alexander Swade

Department of Accounting and Finance

Lancaster University

A thesis submitted in fulfillment of the requirements

for the degree of

Doctor of Philosophy in Finance

October 2023

Abstract

This thesis advances the theory and practice of factor investing by exploring the rich set of developed factors to explain portfolio performances in the equity and multi-asset space. Chapter 1 characterizes the strong performance of equal-weighted (EW) portfolios in relation to their value-weighted counterparts by utilizing various factor models. Unsurprisingly, EW investing comes with a highly significant positive size factor exposure but is also found to benefit from short-term reversal effects while suffering from negative momentum exposure due to its acyclic rebalancing character. Given that EW investing effectively emerges as factor investing in disguise, it seems natural to adopt a direct factor investing approach. To this end, the literature has proposed a multitude of firm characteristics for explaining the cross-section of stock returns, yet Chapter 2 demonstrates only about 15 factors to be relevant for spanning the entire factor zoo from an alpha perspective. Whilst these salient factors change through time, they fall into persistent factor style categories. Further broadening the scope, the thesis moves on to explain the cross-section of asset classes through a macro factor lens. Specifically, Chapter 3 investigates macroeconomic factor allocation based on macro factor-mimicking portfolios that consider style factors and individual asset classes alike. Chapter 4 investigates such macro factor investing over a century of data, demonstrating it to be robust in different economic regimes. Incorporating business cycle-based macro and style factor views in a Black-Litterman fashion we additionally accommodate the notion of factor timing to improve upon a diversified macro factor risk-parity strategy.

The three great essentials to achieve anything worthwhile are, first, hard work; second, stick-to-itiveness; third, common sense.

— *Thomas Edison*

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and acknowledgments.

Alexander Swade

October 2023

Acknowledgements

My doctoral studies have been a remarkable experience which was sometimes challenging but definitely rewarding and worthwhile undertaking. The successful completion of this work would not have been possible without the help and support of many people to whom I am deeply thankful.

First of all, I would like to express my deepest gratitude to my supervisors Prof. Mark Shackleton and Prof. Sandra Nolte (both Lancaster University) as well as my industry supervisor Dr. Harald Lohre (Robeco) for their invaluable feedback, continuous support, and extremely helpful advice throughout the whole journey of my PhD. Especially, their attention to detail paired with deep expertise in the financial sector helped to push my work to a new level. My work has definitely benefited from our fruitful discussions and I have the greatest respect for the passion and enthusiasm they have put in accompanying my doctoral studies.

I am grateful to Lancaster University Management School (LUMS) and the Economic and Social Research Council (UK) for providing me with generous financial support to fund my PhD.

Special thanks go out to my co-authors who helped to leverage my research skills. It was a great pleasure to collaborate with such inspiring researchers and to be challenged throughout the process of making different ideas and thoughts tangible. The resulting research papers are reflected in the four chapters of this cumulative dissertation. As part of the interdisciplinary cooperation between Lancaster University and associated industry partners, these papers are co-authored not only with my supervisors but also with further individuals from associated entities.

As a result, the second chapter, "Compressing the Factor Zoo", is a joint project with Matthias Hanauer (Robeco, TU Munich University) and David Blitz (Robeco). They both contributed to the project by helping develop the research idea as well as providing factual reviews. The third chapter, “Macro factor investing with style”, is a joint project with Scott Hixon and Jay Raol (both Invesco Ltd) who served as contact persons for interested clients (being responsible for the respective capabilities at Invesco).

I am also very thankful for all the friends and colleagues that I have met at LUMS, Invesco, and Robeco during my journey. There are too many names to mention here, but the past four years have certainly been a great experience and joyful episode of my life. Without the many discussions and their support to help with my technical questions, my doctoral studies would have been even harder. Especially, I would like to convey my appreciation and acknowledgments to Filip Bašić, Yves Felker, David Happersberger, Josh Robinson, Carsten Rother, Nikolaos Vasilas, and Shifan Yu for their help, support, and friendship.

Table of Contents

List of Tables

List of Figures

Introduction

Since the introduction of the capital asset pricing model (CAPM)[1], the asset pricing literature has described and analyzed a variety of firm characteristics to explain the cross-section of stock returns. Eventually, the number of characteristics analyzed and factors created resonated in a large set of factors also referred to as 'zoo of factors' (Cochrane, 2011). However, the usage of factors is not limited to explaining the cross-section of stock returns. Many well-documented factors have been identified within and across various alternative asset classes beyond equities. As for the latter, macroeconomic factors have gained increasing interest in the presence of high inflation and other macroeconomic uncertainties.

The aim of this dissertation is to investigate the relevance of factors in explaining key asset allocation concepts, both on a pure equity level but also in a multi-asset context. Specifically, we address three aspects of systematic factor investing. First, we analyze the performance of equal-weighted (EW) portfolios compared to their value-weighted (VW) counterparts. We use classic academic factor models (cf. Fama and French, 1993, 2015 or Hou, Xue, and Zhang, 2015) and their extensions to investigate the long-term evidence for the EW–VW return spread in a broad U.S. equity universe. Unsurprisingly, EW investing comes with a highly significant positive size factor exposure. Given its acyclic rebalancing character, EW investing is also found to benefit from short-term reversal effects while suffering from negative momentum exposure. We also document a pronounced seasonality effect in EW investing that would see outsized returns in January. We revisit these findings in the more investible universe of S&P500 stocks and discuss how to best harvest the embedded factor

[1]See Sharpe (1964), Lintner (1965), Mossin (1966) and Treynor (1961).

premia.

Second, we explore the multitude of firm characteristics that have been deemed relevant factors in explaining the cross-section of equity returns. Yet, the resulting 'factor zoo' can likely be compressed to a few contenders. We aim to pinpoint such factors in chapter 2. Rather than describing the covariance structure of factor returns, we identify factors that best explain the available factor zoo. Our analysis reveals about 15 factors spanning the entire factor zoo, recruiting a few representatives from each factor style. Although the factor styles are persistent, the chosen factor representatives vary over time, illustrating the importance of continuous factor innovation.

Third, we take a macroeconomic stance on systematic investing in chapters 3 and 4. Investors typically face similar macroeconomic risks and opportunities regardless of their individual investment preferences. To best navigate growth and inflation concerns, we propose building macro factor-mimicking portfolios (MFMPs) diversified across asset classes and style factors in chapter 3. We thus focus on the macro factors *Growth*, *Inflation*, and *Defensive* and construct investable MFMPs that mimic their behavior in distinct macroeconomic environments. Our approach also allows for shaping the macroeconomic risk exposure of any given portfolio by applying systematic macro factor completion to effectively address specific economic outcomes.

Consecutively, we extend our understanding of macro factor investing and analyze whether our proposed approach of MFMPs would have been robust over the last century in chapter 4. Using 100 years of global data we analyze their macroeconomic sensitivities and highlight the relevance of navigating time variation in macroeconomic risk premia. Specifically, we adapt the portfolio allocation to align with the identified macro environment as predicted by a forward-looking business cycle model. A Black-Litterman framework is used to thus improve upon a diversified macro factor allocation and to further tap into predictive style factor signals.

This dissertation contributes to the existing literature on systematic factor investing as follows. First, we add to the ongoing debate about the strong performances of EW portfolios, which have been proven to be extremely hard to beat in practice as highlighted by DeMiguel, Garlappi, and Uppal (2009). Our factor approach covering almost six decades of data allows us to analyze the EW-VW spread through different economic periods without limiting the sample to surviving stocks. Second, we add to the debate about the ideal factor model size and also propose a simple yet effective method to identify the relevant alpha contributors in the factor zoo until all remaining alpha sources are exhausted. These alpha sources are of particular interest to practitioners, as they represent the available factor zoo alpha with the minimum number of factors necessary. Third, we advance the macro factor literature by building out robust macro factor mimicking portfolios that incorporate not only asset classes but also style factors. Portfolio allocations based on these MFMPs are tested under different economic scenarios and even further enhanced with an active macro factor timing component in the presence of varying business cycle scenarios.

The remainder of this dissertation is structured as follows. Chapter 1 is based on the research paper titled "Why do equally weighted portfolios beat value-weighted ones?" which has been published in the Journal of Portfolio Management. Therein, we use a set of different factor models to decompose the spread of equal-weighted portfolios and their value-weighted counterparts. Chapter 2 reflects the research paper "Factor zoo (.zip)". We propose an intuitive yet effective methodology to capture about 15 relevant alpha contributors that explain all remaining factor alphas in the zoo of about 150 factors. Chapter 3 originates from the research paper titled "Macro factor investing with style" which has been published in the Journal of Portfolio Management. We construct macro factor mimicking portfolios including style factor and asset class returns to mimic the macro factors *Growth*, *Inflation*, and *Defensive*. Chapter 4 then uses the same approach to analyze macro factor investing over the last 100 years and introduces macro factor views to time the factors in a Black-Litterman fashion. The last chapter concludes this dissertation.

Chapter 1

Why Do Equally Weighted Portfolios Beat Value-Weighted Ones?

This project is joint work with my supervisors Sandra Nolte, Mark Shackleton, and Harald Lohre. It is published in the Journal of Portfolio Management (*Journal of Portfolio Management*, 49 (5), 167–187). We thank David Blitz, Daniel Giamouridis, Amit Goyal, Matthias Hanauer, Pim van Vliet, and participants at the 2022 Frontiers of Factor Investing Conference in Lancaster as well as the 2021 7th International Young Finance Scholars Conference in Oxford for helpful comments and suggestions. This work has been supported by an ESRC NWSSDTP CASE Grant.

1.1 Introduction

The simple approach of equally weighting portfolio constituents is a popular choice of academics and investors to benchmark specific portfolio allocations. Indeed, equal-weighted (EW) strategies prove hard to beat out of sample even when using different optimized asset allocation strategies, see DeMiguel, Garlappi, and Uppal (2009). The success of the EW strategy has piqued many researchers' interest trying to rationalize and exploit its underlying drivers. Many different stock characteristics have been put to the fore, yet, there is no clear evidence which effects drive this outperformance. Given that many of the analyzed stock characteristics seem to only be relevant during specific periods and disappear over time, we take a systematic approach to understand the drivers behind the differences in performance between the EW and value-weighted (VW) portfolios over six decades. Specifically, we analyze the difference of the EW portfolio and its market capitalization weighted counterpart (also referred to as the VW portfolio) to differentiate between persistent and transitory performance components. We thus analyse multiple setups ranging from single to multi-factor models utilizing well-known factors. To investigate the practicality of our findings, we do not only focus on the broad CRSP universe but also the large-cap S&P 500 universe. Despite some sample-specific differences, we find that the vast majority of systematic effects carry over.

Our work is related to recent research exploring the equal-weighting scheme and its performance, e.g., Malladi and Fabozzi (2017), Pae and Sabbaghi (2015), and Plyakha, Uppal, and Vilkov (2021), as well as size factor related literature like Asness et al. (2018) and Blitz and Hanauer (2020). Whilst previous research focuses on stock-specific characteristics to explain the different return profiles of the EW and VW portfolios, we take a closer look at the contribution of systematic factors to the EW–VW spread over time and across different factor models. Many of the previous papers pinpoint the different effects in terms of one or two components which prove consistent through the respective sample. The most obvious ones are the size tilt of the EW portfolio towards small caps as well as the rebalancing effects that derive from the necessity to maintain equal portfolio weights. We confirm that the size factor is the most significant driver of the performance of the difference between the EW and VW portfolio, yet, we also highlight the impact of other factors beyond size. Notably, factors such as momentum, profitability, short-term reversals, or low volatility also help to increase the explained return variation within models during certain time periods.

We contribute to the literature in several ways. First, we analyze the impact of equal-weighting over six decades for the broad CRSP equity universe, decomposing the long-short EW–VW return spread into its systematic components. Second, we test a variety of factor models ranging from a single index model to well-known multi-factor models such as the one proposed by Fama and French (2015) to further decompose the EW–VW spread. We confirm size as the prevailing factor component but also emphasize the relevance of other factors to explain the time-varying magnitude of the difference between the EW and VW portfolio. Third, we document the close relation of the EW–VW spread and the small minus big (SMB) size factor, resulting in an easy to implement alternative to small cap funds to harvest the size premium.

The remainder of this paper is structured as follows. Section 1.2 reviews the literature on EW investing and describes the different factor model frameworks, capturing the single index model (SIM) and multi-factor models. Section 1.3 documents the historical outperformance of the EW portfolio relative to its VW counterpart in the CRSP and S&P 500 universes and associates it with multiple systematic components. Further analysis highlights the seasonality in the EW–VW spread. Section 1.4 investigates the possibility to harvest the size premium by investing in the EW–VW spread relative to purchasing small cap funds. Section 1.5 concludes.

1.2 Setting the theoretical foundation

1.2.1 The virtue of equally weighting

Analyzing the performance of EW (or $1/N$) portfolios has garnered considerable interest from academics and practitioners alike. It is systematic and easy to implement because all N portfolio constituents are assigned the same weight, $w_i = \frac{1}{N}$ for $i = 1, ..., N$. It is though an active strategy since the portfolio requires rebalancing to maintain equal weights over time. At each rebalancing date, it sells winners and buys losers and is thus considered a mean-reversion, contrarian strategy yielding concave payoffs equivalent to selling portfolio insurance (Perold and Sharpe, 1988). The deterministic weighting scheme does not require any expected return or variance input and intrinsically enables diversification. Therein, a naive investor is only reliant on the average correlation coefficient to determine acceptable risk-return trade-offs (De Wit, 1998). These features make the EW portfolio a strong contender compared to different allocation schemes as highlighted by DeMiguel, Garlappi, and Uppal (2009). In contrast, the VW portfolio is a passive buy and hold strategy which reflects market drifts. Historically, investing in EW portfolios was rewarded with a premium compared to investing in the corresponding VW portfolios. Kaiser and Peter (2022) find improved returns when following a rotation strategy between the VW and EW portfolio based on lagged one-month market returns.

Several explanations have been proposed to rationalize the exceptional performance of EW portfolios relative to alternative allocation methods. Plyakha, Uppal, and Vilkov (2021) document a monotonic relation[1] between the EW–VW return spread and size, price, liquidity, and idiosyncratic volatility factors. Put differently, the higher the stocks' characteristics in the sampled portfolio, the larger the resulting EW–VW spread. They link the higher returns of the EW portfolios to systematically higher exposures to market, size and value factors; still, EW portfolios exhibit significantly positive four-factor alpha in the sense of Fama and French (1993) and Carhart (1997) models which they rationalize with the need for frequent rebalancing to maintain equal portfolio weights.

Naturally, EW investing can also be related to the rebalancing literature which suggests that there are benefits to the mere act of rebalancing, labeled *diversification return* (Erb and Harvey, 2006), *volatility return* (Willenbrock, 2011) or *rebalancing premium* (Bouchey et al., 2012). In this vein, a proportion of the difference between the EW and VW returns might just be related to the *rebalancing return* defined as the difference in growth rates of the rebalanced versus the buy-and-hold portfolio. Hallerbach (2014) analytically shows that the rebalancing return emerges as the difference between a volatility return and a dispersion discount. Since both components are strictly positive, it is not a given whether the rebalancing return is

[1] Based on the methodology of Patton and Timmermann (2010).

positive or negative. To investigate, one could split the return difference between the EW and VW portfolios into a rebalancing component return, as well as a component related to the difference in weighting. However, in order to cleanly split the rebalancing effects, one must have the same constituents over the full sample period, because any change of constituents involves rebalancing, which interferes with the buy and hold strategies. Against this backdrop, Maeso and Martellini (2020) analyze rebalancing returns for a constant sample of surviving S&P 500 stocks and identify a sizable annual premium of the rebalanced strategy over its buy-and-hold counterpart. Malladi and Fabozzi (2017) develop a two-period, two-asset model in which the difference in weighting and the rebalancing effect are the two sources for outperformance. Their empirical results include randomized prices which also eliminate the need to rebalance because of changes in constituents. Given the said caveat of using a constant sample of companies, we rather build our analysis on a more realistic setup that considers all investible companies at any point in time. Therefore, we cannot gauge potential rebalancing premia; yet, we are optimistic to rationalize the EW-VW spread in terms of different factor premia that will most likely play an important role in rationalizing rebalancing premia as well.

Another strand of the literature advocates market-related rationales to explain the differences between EW and VW portfolio performance: The *noisy market hypothesis*[2] expands the *efficient market hypothesis* by constructing a theory where securities are not always priced at their fair values but are over- or undervalued because of market inefficiencies. To deviate from market capitalization-weighted indexes, Arnott, Hsu, and Moore (2005) analyze several alternative fundamental measures besides firms' market capitalization. The resulting indexing schemes are referred to as *fundamental indexing* and the EW index can be considered an alternative in that regard. The initial claim of Arnott, Hsu, and Moore (2005) that VW portfolios are tilted towards over-valued stocks and under-represent value firms has been debated since, e.g., by Perold (2007) or Kaplan (2008). Despite the debate of its validity, the fundamental indexing framework implies that the difference in performance between the EW and VW portfolios is due to mispricing of over-weighted firms in the VW portfolio. Hence, investing in an EW portfolio instead can be interpreted as investing in the

[2]The term *noisy market hypothesis* was introduced by Siegel (2006).

passive VW portfolio plus an additional overlay similar to the SMB and HML factors by Fama and French (1993).

1.2.2 Factor models and EW investing

A SIM approach

We start exploring systematic effects of the EW–VW return differences by invoking the simplest factor model – the single index model (SIM) also known as the capital asset pricing model (CAPM).[3] Therein, the return of stock i is explained by its market beta, β_i (against the market return, r_M, with expected return, $\overline{r}_M$, and variance, σ_M^2), its alpha, α_i, and some zero-mean idiosyncratic risk ε_i. Empirically, this model aims for alphas to be statistically indistinguishable from zero to have high explanatory power from the systematic factor(s) of the model. The idiosyncratic risks are assumed to be stock-specific with an individual but independent magnitude σ_i^2 all equal to a common idiosyncratic variance, σ_I^2. In this case, the return, r_i, of stock i, its expectation and variance are given by

$$\begin{aligned} r_i &= (\alpha_i) + \beta_i r_M + \varepsilon_i, \\ E\left[r_i\right] &= \beta_i E\left[r_M\right] = \beta_i \overline{r}_M, \\ Var\left[r_i\right] &= \beta_i^2 \sigma_M^2 + \sigma_I^2. \end{aligned} \tag{1.1}$$

The returns r_i and r_M are in excess of the risk-free rate. Note that all parameters in the model are time-varying, but we omit time indices for readability.

We denote the value-based weights associated with the market capitalization of stock i as w_i^{mcap} and the VW portfolio return as r_{VW} with expected return $\overline{r}_{VW}$ matching that of the market:

$$r_{VW} = \sum_{i=1}^N w_i^{mcap} \left(\beta_i r_M + \varepsilon_i\right) = r_M \sum_{i=1}^N w_i^{mcap} \beta_i + \sum_{i=1}^N w_i^{mcap} \varepsilon_i. \tag{1.2}$$

From equation (1.2) we can infer that $\sum_{i=1}^N w_i^{mcap} \beta_i = 1$, which holds by definition.[4] As for

[3]See Sharpe (1964), Lintner (1965), Mossin (1966) and Treynor (1961).

[4]The single index model market return has a beta of 1 by definition. In our case, the market return is defined as the value weighted portfolio return.

the idiosyncratic risk, $\sum_{i=1}^{N} w_i^{mcap}\varepsilon_i$ goes to 0 for large N, assuming the stocks' idiosyncratic risks to be unbiased and independent of weights such that the weighted average error will be zero. The variance of this value weighted portfolio equals the variance of the market, i.e., $\sigma_{VW}^2 = \sigma_M^2$ as $\sigma_I^2/N \to 0$ for $N \to \infty$.

The EW portfolio has fixed weights $w_i = 1/N$ for all i. The return of the EW portfolio is labeled r_{EW}, and its expected return $\overline{r}_{EW}$ emerges as the sample average of beta times the expected market return $\overline{r}_M$.

$$\begin{aligned} r_{EW} &= \sum_{i=1}^{N} \tfrac{1}{N}\left(\beta_i r_M + \varepsilon_i\right) = r_M \tfrac{1}{N}\sum_{i=1}^{N}\beta_i + \tfrac{1}{N}\sum_{i=1}^{N}\varepsilon_i, \\ \overline{r}_{EW} &= E\left[r_M \tfrac{1}{N}\sum_{i=1}^{N}\beta_i\right] + E\left[\tfrac{1}{N}\sum_{i=1}^{N}\varepsilon_i\right] = \tfrac{1}{N}\sum_{i=1}^{N}\beta_i \overline{r}_M = \overline{\beta}\,\overline{r}_M, \end{aligned} \tag{1.3}$$

where $\overline{\beta} = \frac{1}{N}\Sigma_{i=1}^{N}\beta_i$ is the sample average of beta which may deviate from unity. The variance of the EW portfolio is given by $\sigma_{EW}^2 = \overline{\beta}^2\sigma_M^2$, that is, the EW portfolio is locally a $\overline{\beta}$ multiplier of the VW passive portfolio where $\overline{\beta}$ is scaling expected return and market volatility.

Using (1.2) and (1.3) we can express the difference in performance between EW and VW portfolios. Equation (1.4) thus suggests why and when the EW portfolio outperforms the VW one: The average constituent's beta, $\overline{\beta}$, intensifies the market return if $\overline{\beta} > 1$ and abates it if $\overline{\beta} < 1$. Hence, if $\overline{\beta} > 1$, the EW portfolio outperforms the VW one for positive r_M and vice versa, the relation is reversed for $\overline{\beta} < 1$.

$$r_{EW} - r_{VW} = \overline{\beta} r_M - r_M = \left(\overline{\beta} - 1\right) r_M. \tag{1.4}$$

By construction, the expected return correlation of EW and VW portfolios is 1. Empirically, the correlation of returns is less than one, given that the error term assumptions do not hold perfectly. Also, the drift in weights due to non-continuous rebalancing will break this relation occasionally.

Multi-factor models

The market factor in the single index model might not capture all systematic risk sources, so that further linear models such as Ross's (1976) arbitrage pricing theory (APT) have been developed. Under APT, the returns $\boldsymbol{r} \in \mathbb{R}^{N\times 1}$ of N risky assets follow a factor intensity structure given by:

$$\boldsymbol{r} = \boldsymbol{B} \cdot \boldsymbol{f} + \boldsymbol{\varepsilon}, \tag{1.5}$$

where $\boldsymbol{f} \in \mathbb{R}^{K\times 1}$ represents the returns of K factors with respective factor loadings $\boldsymbol{B} \in \mathbb{R}^{N\times K}$ and asset-specific idiosyncratic risks $\boldsymbol{\varepsilon} \in \mathbb{R}^{N\times 1}$, which have zero mean and are assumed to be uncorrelated across assets and factors. The expected asset returns can be expressed in terms of factor sensitivities, so that:

$$\mathbb{E}\left(\boldsymbol{r}\right) = \boldsymbol{r_f} + \boldsymbol{B} \cdot \boldsymbol{r_p}, \tag{1.6}$$

with $\boldsymbol{r_f} \in \mathbb{R}^{N\times 1}$ denoting the risk-free rate, and $\boldsymbol{r_p} \in \mathbb{R}^{K\times 1}$ denoting the risk premia associated with the corresponding systematic factors. Several factor models follow this paradigm, e.g., Fama and French (1993), Carhart (1997), Hou, Xue, and Zhang (2015), and Fama and French (2015). Our choice of factors is described in the next chapter.

1.3 Analyzing the EW-VW spread

1.3.1 The historical outperformance of the EW portfolio

Empirical setup

To empirically investigate the equal-weighted portfolio, we construct the market-weighted portfolio as well as its equally weighted counterpart for a broad US equity universe as well as the S&P 500 universe. Our full sample period spans from July 1963 to December 2021, and we use monthly data from CRSP and Compustat covering stocks traded on the NYSE, AMEX, and NASDAQ with share codes 10 or 11. Figure 1.1 reveals the dominance of the EW portfolio over the long-term, depicting the performance of the two portfolios based on the full CRSP universe (EW) and the S&P 500 (SPW) constituents only. In the CRSP universe, EW gives positive returns of 14.9% p.a. at a volatility of 20.6%; in the S&P 500 universe we obtain 13.6% p.a. at 17.1% volatility. Both VW portfolios show similar annual returns (11.4%) at some 15% volatility over the same period. Note that the EW portfolios

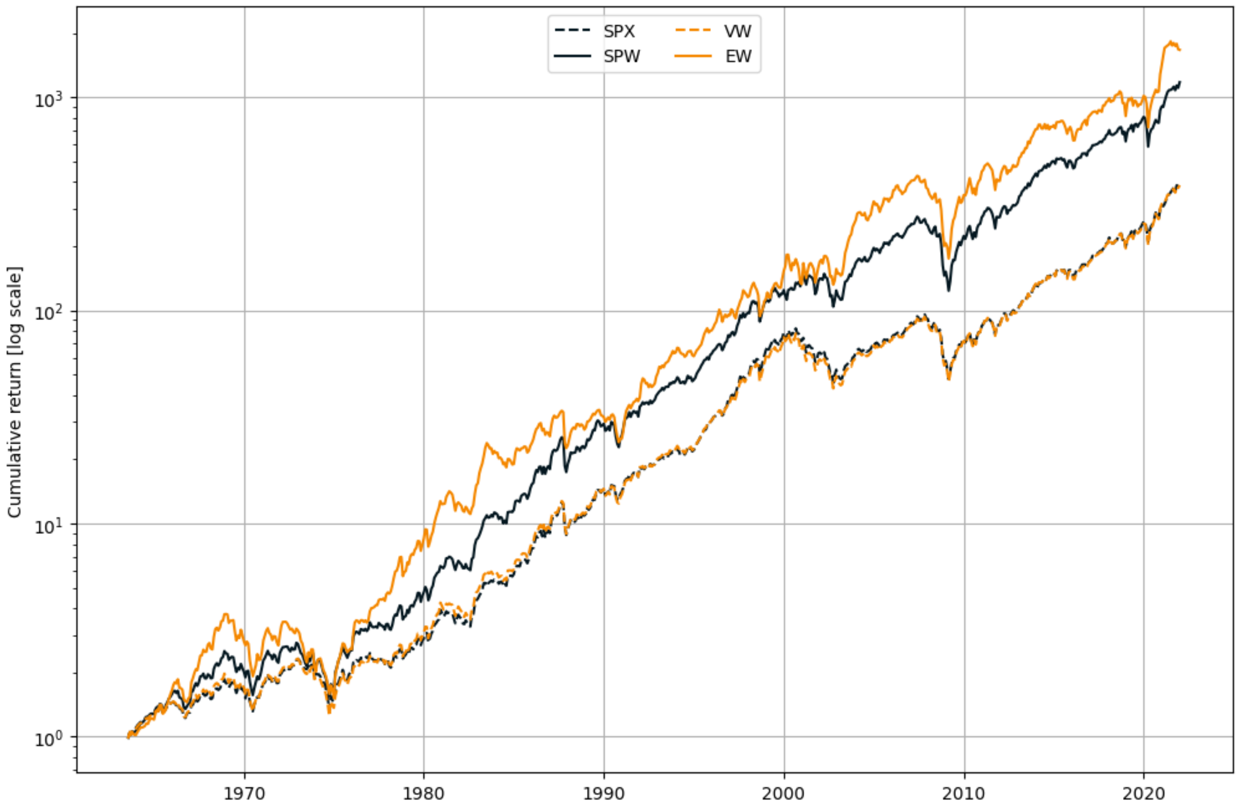


Figure 1.1: Cumulative Performance EW and VW Portfolios. This figure depicts the performance of value- and equal-weighted portfolios based on the S&P 500 index (SPX, SPW), and all traded stocks on the NYSE, AMEX, and NASDAQ with share codes 10 and 11 (VW, EW). The sample period is July 31, 1963–December 31, 2021.

outperform their VW counterparts in 33 (37) out of the 59 years, an outperformance which comes at higher risks in terms of higher volatility and more severe drawdowns.

Table 1.1 presents the performance differences between EW and VW portfolios for the CRSP (Panel A) and S&P 500 samples (Panel B) across different time periods. These subperiods are the pre-publication (July 1963–December 1983) and post-publication (January 1984–December 1999) periods referring to the first size effect publications[5], the period before (January 2000–December 2009) the global financial crisis (GFC), as well as the time after it (January 2010–December 2021). The splits in these four subperiods resonate with the strong and weak performance periods of the EW portfolio compared to its VW counterpart: Prior publication as well as prior to the GFC, the EW portfolios outperformed across both

[5]See Banz (1981) or Reinganum (1981)

universes yielding annual excess returns of 7.5% and 10.1% (4.0% and 7.2% for the S&P 500 universe) over the VW portfolios. Despite higher volatility and maximum drawdown (MDD) figures, the EW portfolios surpassed the VW ones on a risk-adjusted basis yielding Sharpe ratio (SR) differences of 0.08 and 0.13 (0.05 and 0.11 for S&P 500). In contrast, EW portfolios did considerably worse during the other two subperiods. In the post-publication phase, the EW portfolios underperformed by -3.1% p.a. in the CRSP and -1.5% p.a. in the S&P 500 sample, whereas they went flat after the GFC, showing annual return differences of -0.1% and 0.0%. Yet, the EW portfolios come with higher risks during all periods resulting in risk-adjusted underperformances in the latter periods as well. These differences over time are most likely driven by the following phenomena: After the discovery of the size effect by Banz (1981) and Reinganum (1981), academics as well as practitioners started to account for this anomaly resulting in weaker performances of the EW portfolios thereafter, which are mechanically linked to small firms' performances. EW portfolios seem to recover well just after big market corrections like the burst of the dotcom bubble in 2001 and the GFC in 2008. Under-weighting large caps helped mitigating the extreme drawdowns but also participating in size effects during the recovery, ultimately boosting the spread in performance. In addition, EW portfolios lag in times of monotonic market trends and dominance of large caps, such as the FAANG[6] stocks rally during the 2010s.

Table 1.1 also splits the returns of the VW and EW portfolios into the month of January versus the period from February through December (non-January), addressing the seasonality of the size effect as highlighted by, e.g., Keim (1983) and Roll (1983). Indeed, we observe considerably higher annualized returns for the month of January compared to non-January months in both universes and across portfolios. This effect becomes even stronger for the EW portfolios resulting in 2.5–6 times higher returns over the non-January months. Finally, the last two columns of Panels A and B of Table 1.1 report the average monthly return difference between the EW and VW portfolios for all subperiods as well as t-statistics for testing the hypothesis of returns being zero. Across both Panels we identify significant outperformances of the EW portfolios for the full sample period as well as the pre-publication

[6]This acronym refers to the five best-performing U.S. technology firms: Facebook, Amazon, Apple, Netflix and Alphabet (Google).

Table 1.1: The Effect of Equal-weighting across Sample Periods and Universes

Sample	*Years*	*Ret p.a.*		*Std p.a.*		*SR*		*MDD*		*Mcap*	*Const*	*EW–VW*	
		VW	*EW*	*VW*	*EW*	*VW*	*EW*	*VW*	*EW*			*ret*	*t-stat*
Panel A: CRSP													
Full sample	1963–2021	11.4	14.9	15.3	20.6	0.13	0.15	-22.6	-28.2	2.1	4,349	0.29	2.42
January		17.3	65.4	17.5	26.0	0.21	0.68	-8.1	-9.5	2.0	4,372	4.01	6.57
Non-January		10.9	10.3	15.1	19.6	0.12	0.09	-22.6	-28.2	2.1	4,347	−0.05	−0.42
Pre-publication (Expansion)	1963–1983	9.8	17.3	15.3	21.6	0.06	0.14	-12.2	-18.8	0.3	3,172	0.63	3.01
Post-publication (Downfall)	1984–1999	17.1	14.0	15.1	17.7	0.22	0.14	-22.6	-28.2	0.8	6,003	−0.26	−1.25
Pre-GFC (Recovery)	2000–2009	1.0	11.1	16.5	24.1	−0.03	0.10	-17.1	-21.2	2.7	4,959	0.84	2.49
Post-GFC (Stagnation)	2010–2021	15.1	15.0	14.4	19.6	0.29	0.22	-13.2	-22.4	6.4	3,648	−0.01	−0.03
Panel B: S&P 500													
Full sample	1963–2021	11.4	13.6	14.8	17.1	0.14	0.16	-21.6	-25.6	14.2	500	0.19	2.99
January		15.2	27.6	17.1	20.6	0.18	0.33	-8.3	-7.8	13.8	500	1.03	3.40
Non-January		11.0	12.4	14.6	16.8	0.13	0.14	-21.6	-25.6	14.3	500	0.11	1.80
Pre-publication (Expansion)	1963–1983	9.3	13.3	14.4	17.4	0.06	0.11	-11.8	-15.1	1.1	499	0.33	2.80
Post-publication (Downfall)	1984–1999	18.0	16.5	14.9	16.3	0.24	0.19	-21.6	-25.6	7.3	500	−0.12	−1.23
Pre-GFC (Recovery)	2000–2009	0.6	7.8	16.1	19.4	−0.04	0.07	-16.7	-20.8	21.5	500	0.60	3.35
Post-GFC (Stagnation)	2010–2021	15.2	15.2	13.8	15.8	0.31	0.27	-12.2	-18.8	39.8	503	0.00	−0.01

Panel C: Single Index model results			CRSP					S&P 500				
Sample	*Years*	*Portfolio*	α	t(α)	β	t(β)	R^2	α	t(α)	β	t(β)	R^2
Full sample	1963–2021	VW	0.00	−1.25	1.00	−5.28	1.00	0.02	0.87	0.95	−8.75	0.98
		EW	0.20	1.67	1.15	5.77	0.73	0.14	2.56	1.07	5.55	0.92
		EW–VW	0.20	1.68	0.15	5.85	0.05	0.12	2.00	0.11	8.44	0.09
Pre-publication (Expansion)	1963–1983	EW–VW	0.56	2.82	0.23	5.19	0.10	0.29	2.61	0.15	6.13	0.13
Post-publication (Downfall)	1984–1999	EW–VW	−0.24	−1.09	−0.03	−0.56	0.00	-0.19	−1.96	0.07	3.39	0.06
Pre-GFC (Recovery)	2000–2009	EW–VW	0.88	2.74	0.26	3.90	0.11	0.62	3.63	0.13	3.71	0.10
Post-GFC (Stagnation)	2010–2021	EW–VW	−0.27	−1.25	0.22	4.29	0.11	-0.14	−1.46	0.11	5.09	0.15

This table reports key performance statistics of the VW and EW portfolios over time. Panels A and B focus on the CRSP and S&P 500 sample, respectively. Return, volatility and 1-month maximum drawdown (MDD) are in percentage terms. Average market capitalization is in billion USD. The last two columns report the monthly average return difference between the EW and VW portfolios, as well as its t-statistic. Panel C reports Single Index Model results for both universes. α values are reported in percentage points per month. α t-stats are reported against the hypothesis of $\alpha = 0$. β t-stats are reported against the hypothesis of $\beta = 1$ for the two portfolios (EW, VW) and $\beta = 0$ for the EW–VW spread. The full sample period is July 31, 1963–December 31, 2021. All other sub-samples start in January and end in December of the reported years (except for the Pre-publication period, which starts end of July). January and Non-January (February–December) statistics are reported for the full sample period.

and pre-GFC subperiods of 29 to 84 bps per month (with t-stats ranging from 2.42 to 3.35). This outperformance is more pronounced during the month of January with 401 bps (t-stat 6.57) for the CRSP sample and 103 bps (t-stat 3.40) for the S&P 500 sample.

Thus, the effects of equal-weighting compared to value-weighting are significant for both universes despite considerable differences in universes: Whilst the S&P 500 has 500 constituents by definition[7], the CRSP universe increases from 3,172 in the pre-publication phase[8] to about 6,000 in the post-publication period before shrinking back to 3,648 over the last decade. This development highlights the impact of extremely small companies, which is also reflected in the average market capitalization (Mcap) of the portfolio constituents. While the average Mcap in the whole universe increases from 0.3 to 6.4 billion USD, the average S&P 500 firm is around 5 times bigger than the average firm in the CRSP universe peaking at a factor of more than 9 during the 1990s. These initial observations call for systematically analyzing factors driving the performance difference, which we investigate next.

A first glance using the single index model

To begin with, we estimate the SIM using the market factor MKT in excess of the risk-free rate as provided by Fama and French (1993). To account for the time-varying characteristic of individual stocks' beta as well as its short-term persistence we run regressions based on the different subperiods. Panel 3 of Table 1.1 depicts the estimated model parameters α and β as well as their t-statistics and the overall R^2 for the six CRSP and S&P 500 portfolios over the full sample period. These are the value-weighted one, its equal-weighted counterpart as well as the spread of EW–VW for each sample. The subsequent rows depict the parameter estimates for the EW–VW spreads in the highlighted subsamples.

The full sample CRSP VW portfolio shows the expected model parameters with α being statistically indistinguishable from 0 while β is equal to 1. As expected, the resulting R^2 is 100%. For the S&P 500 universe we observe similar results; yet, the estimated R^2 of 98% for the VW portfolio indicates a slight variation from the market portfolio by missing out a

[7]Note that the exact number of available stocks based on the CRSP database might slightly deviate from 500 due to information lags between CRSP's and Standard and Poor's listing dates. Additionally, the S&P 500 includes several companies with two share classes increasing the total count, e.g., Alphabet's Class A (GOOGL) and Class C (GOOG) shares.

[8]The CRSP universe consists of 1,741 stocks at the start of the sample period in the early 1960s.

considerable fraction of small caps. With β being significantly below 1 (t-stat -8.75; tested against hypothesis of H_0: $\beta = 1$) highlights the reduced market sensitivity. Conversely, the EW portfolios have significant betas of 1.15 (t-stat 5.77, H_0: $\beta = 1$) for the CRSP universe and 1.07 (t-stat 5.55, H_0: $\beta = 1$) for the S&P 500 universe. The models' α increases to 0.20 (t-stat 1.67) and 0.14 (t-stat 2.56). At the same time, the unexplained return variation increases compared to the VW portfolios resulting in R^2 of 0.73 and 0.92, respectively. Regressing the EW–VW spread on MKT for the CRSP and S&P 500 universes gives significantly positive betas, however the SIM does only explain 5% and 9% of the respective return variations.

Analysing the EW–VW spread returns for different periods we report time-varying estimates. During the pre-publication and pre-GFC periods the spreads' alphas are significantly positive (t-stats ranging from 2.61 to 3.63) whilst the spreads' betas are significant with coefficients of 0.23 and 0.26 (t-stats 5.19 and 3.90) for the CRSP universe, and 0.15 and 0.13 (t-stats 6.13 and 3.71) for the S&P 500 universe, respectively. These results indicate that the spread returns have benefited from the average (1/N) firm's beta $\overline{\beta}$ being greater than 1 as well as some further idiosyncratic effects. Yet, the SIM merely explains between 10% to 13% of return variation. During the post-publication and post-GFC periods however, alphas are negative and statistically insignificant whilst spreads' betas are closer to 0 (indicating that $\overline{\beta}$ is close to 1).

Given the above analysis, the SIM helps to explain the VW and EW portfolios' performances but fails to explain the differential performance. The latter effect is more pronounced for the CRSP universe where large caps are being extremely under-weighted with weights around 0.02% in each stock whereas the EW version in the S&P 500 universe applies portfolio weights of 0.2%. Naturally, this outcome calls for additional systematic factors to help explain the EW-VW spread.

1.3.2 The EW–VW spread through a multi-factor lens

Factor set

To investigate the systematic drivers of the EW–VW spread, we focus on a set of common factors used among academics as well as practitioners. Specifically, we analyze the relevance of the size (SMB), value (HML), profitability (RMW), investment (CMA), momentum (WML), and short-term reversal (STR) factors provided by Kenneth R. French[9]. We also add the volatility (VOL) factor by Van Vliet and De Koning (2017)[10] to account for the low-risk anomaly. To test for robustness, we also use the quality-minus-junk (QMJ) factor of Asness, Frazzini, and Pedersen (2019)[11], as well as market equity (ME), investment to asset (IA), return on equity (ROE), and expected growth (EG) from the q-factor database[12] as introduced by Hou et al. (2021) and Hou, Xue, and Zhang (2015).

Our prior is that the EW–VW spread benefits from size as well as short-term reversal effects because of over-weighting small firms and a contrarian rebalancing style. In a similar vein, the spread should be negatively correlated to momentum, which thrives if winners continue to perform well, as well as volatility due to its increased risk. Based on the findings of Asness et al. (2018) and Blitz and Hanauer (2020) we expect the EW-VW spread to also be negatively correlated to the quality factor because of its close link to the size factor and no natural control for junk firms among the small stocks. Figure 1.2 depicts the correlation structure of the factors used in our analysis.

First, we note that the EW–VW spread in the CRSP universe is highly positively correlated with the S&P 500 spread (0.63) indicating that similar effects drive the performance of both spread portfolios. Second, the size factor SMB is highly positively correlated with these spreads (0.87 and 0.63), highlighting their exposure to small firms. Size shows some negative correlation with the profitability factor RMW (-0.35) and is highly positively correlated with the alternative size definition ME (0.97).

At the same time, the momentum factor WML is negatively correlated with the EW–VW spreads (-0.26 and -0.43), as is profitability (-0.40 and -0.12). In contrast, the short-term

[9]http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html
[10]https://www.paradoxinvesting.com/data/
[11]https://www.aqr.com/Insights/Datasets
[12]http://global-q.org/index.html

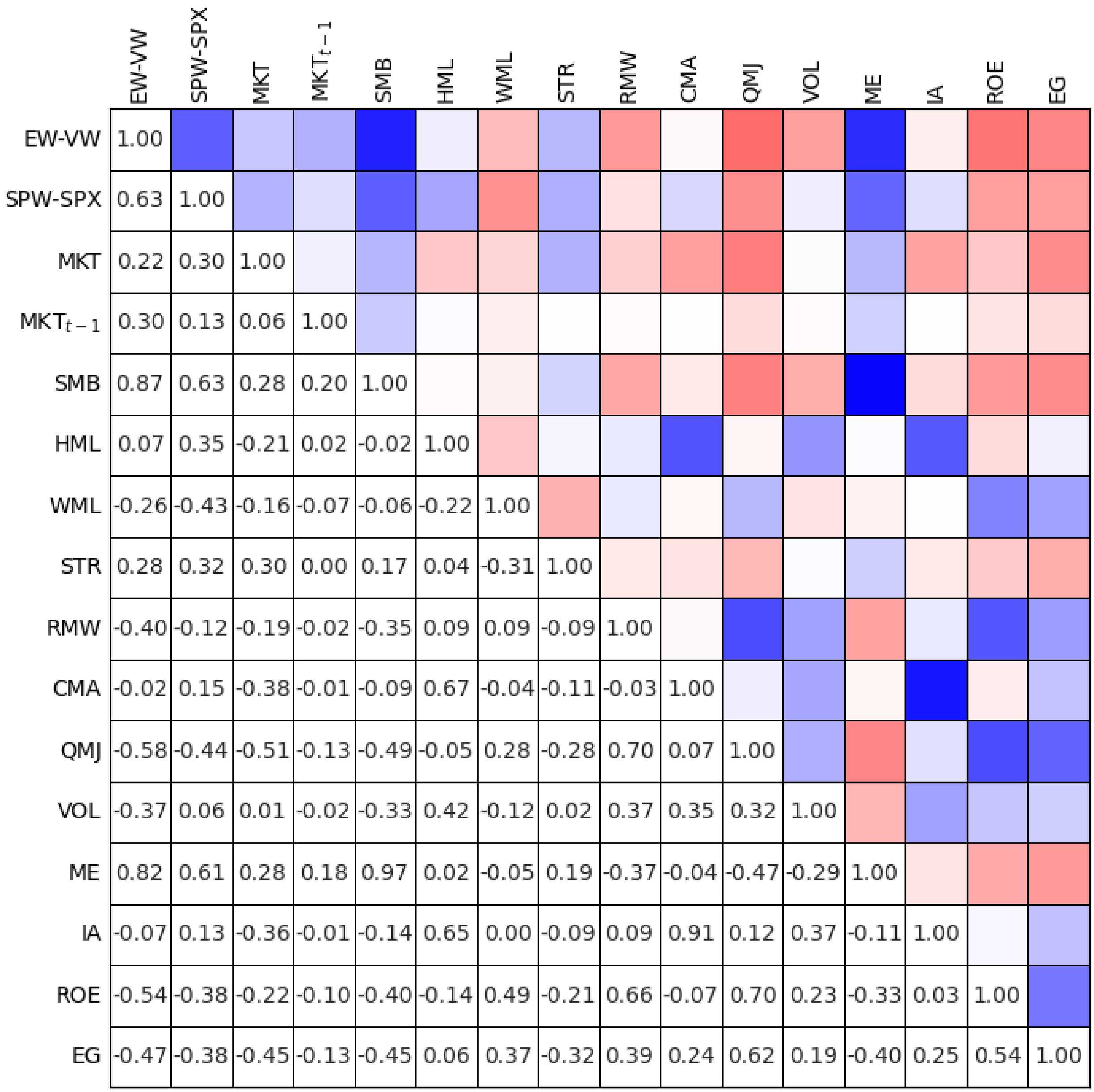


Figure 1.2: Correlation Matrix for Multi-Factor Universe. This figure depicts the correlation structure of the multi-factor universe, building on monthly data for the full sample period July 31, 1963–December 31, 2021, except for the q-factors (ME, IA, ROE, EG), which start on January 31, 1967. Colors range from dark red (correlation of -1) to dark blue (correlation of 1).

reversal factor is positively correlated with the spread, yielding correlation coefficients of 0.28 and 0.32 for the CRSP and S&P 500 universes, respectively. The volatility factor shows negative correlations with the EW–VW spread of the CRSP sample (-0.37). The additional factors of the q-factor model show high correlations of 0.91 (IA vs. CMA) and 0.66 (ROE vs. RMW) with their Fama French counterparts as well as high correlations with QMJ (0.70 for ROE, 0.62 for EG). Most of the other factors seem to be uncorrelated.

Such eyeballing of the underlying correlation structures confirms the size and short-term reversal tilt of the EW-VW spreads as well as negative momentum and quality exposures. Narrowing down the sample to the 500 largest stocks exacerbates the negative momentum exposure whilst attenuating the negative quality exposure due to higher concentration of blue-chip firms amongst large caps.

Multi-factor regressions

In this section, we extend the SIM and investigate a variety of multi-factor models, seeking to further rationalize the EW–VW spread. Table 1.2 depicts the corresponding regression coefficients as well as t-stats for various common factor models[13] for the full sample period from July 1963 to December 2021.

Based on our initial observations indicating a close link between size and the EW–VW spread as well as findings of Asness et al. (2018) and Blitz and Hanauer (2020), we include one month lagged market returns (MKT_{t-1}) in our models to account for non-synchronous trading of small stocks. Thus, we present the single index model results with additional lagged market returns in the first (eighth) column for our full period analysis of the two samples. Documenting significantly positive coefficients of 0.14 and 0.11 (t-stats 5.63 and 8.29) for the market factor indeed suggests the average stock's market beta $\overline{\beta}$ to be greater than one, i.e., $\overline{\beta}$ would be estimated as 1.14 (1.11) in the sample period. Highly significant positive loadings of the lagged market return in the CRSP universe (t-stat 8.30) and still significant ones for the S&P 500 (t-stat 3.21) indicate illiquidity effects amongst the smaller stocks in both samples. However, the adjusted R^2s are still small for both samples (13% and 10%) leaving a lot of unexplained variation in the EW–VW returns.

Second, we learn that regressing the EW–VW spread univariately on SMB gives an adjusted R^2 of 75% with a highly significant t-stat of 45.79, suggesting the spread to be

[13]Note that we report factor exposures based on value-weighted factors. By and large, we confirm our findings when using equal-weighted factor versions. Specifically, EW factors are slightly more relevant explaining the EW–VW spread in the CRSP universe and explain slightly less return variation in the S&P 500 sample. These intuitive results underline the dominance of the largest stocks in a VW setup which become proportionally less significant in EW factor portfolios.

Table 1.2: Factor Regressions of EW–VW Spread Returns

$$EW - VW_t = \alpha + \alpha_{Jan} + \beta_1 F_t^1 + ... + \beta_K F_t^K + \epsilon_t \quad for\ K\ factors$$

	CRSP							*S&P 500*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
α	0.08	0.08	0.10	0.31	0.24	0.17	0.01	0.10	0.11	-0.03	0.08	0.02	0.04	0.03
	(0.75)	(1.38)	(1.64)	(4.46)	(4.21)	(3.32)	(0.16)	(1.63)	(2.23)	(-0.81)	(1.39)	(0.44)	(0.92)	(0.74)
α_{Jan}							1.93							0.07
							(11.85)							(0.53)
MKT	0.14		-0.03	-0.05	-0.09	-0.03	-0.03	0.11		0.09	0.07	0.08	0.06	0.06
	(5.63)		(-1.82)	(-3.37)	(-6.40)	(-2.36)	(-2.52)	(8.29)		(9.10)	(6.03)	(7.11)	(6.23)	(6.23)
MKT_{t-1}	0.21		0.10	0.10	0.09	0.10	0.09	0.04		-0.00	0.00	-0.00	-0.01	-0.01
	(8.30)		(7.80)	(7.88)	(7.41)	(8.94)	(9.50)	(3.21)		(-0.45)	(0.19)	(-0.27)	(-1.24)	(-1.25)
SMB		0.91	0.84		0.78	0.79	0.78		0.34	0.33		0.31	0.35	0.35
		(45.79)	(40.67)		(39.18)	(42.69)	(46.34)		(21.53)	(22.31)		(19.54)	(24.95)	(24.90)
HML			0.13		0.05	0.08	0.06			0.21		0.23	0.13	0.13
			(4.87)		(2.45)	(3.54)	(2.95)			(10.90)		(15.68)	(7.31)	(7.25)
WML						-0.14	-0.13						-0.12	-0.12
						(-11.05)	(-11.36)						(-12.72)	(-12.65)
RMW			-0.19			-0.09	-0.06			0.08			0.10	0.10
			(-7.08)			(-3.46)	(-2.76)			(4.16)			(5.05)	(5.07)
CMA			-0.07			0.06	0.05			0.05			0.08	0.08
			(-1.83)			(1.74)	(1.58)			(1.63)			(2.99)	(2.97)
STR						0.11	0.09						0.04	0.04
						(6.48)	(5.59)						(3.18)	(3.09)
QMJ					-0.36							-0.03		
					(-12.14)							(-1.09)		
VOL						-0.16	-0.13						0.03	0.03
						(-8.66)	(-7.94)						(1.98)	(2.02)
ME				0.74							0.28			
				(34.15)							(16.04)			
IA				0.01							0.24			
				(0.27)							(9.24)			
ROE				-0.35							-0.09			
				(-12.69)							(-3.91)			
EG				-0.10							-0.07			
				(-2.52)							(-2.28)			
Adj. R^2	0.13	0.75	0.79	0.78	0.81	0.85	0.87	0.10	0.40	0.58	0.48	0.57	0.68	0.68
Obs.	701	702	701	660	701	701	701	701	702	701	660	701	701	701

This table presents the factor sensitivities of the EW–VW spread for the CRSP and S&P 500 samples. Excess market return (MKT), its lagged version (MKT_{t-1}), size (SMB), value (HML), profitability (RMW), and investment (CMA) refer to the factors provided by Fama and French (2015). Momentum (WML) and short-term reversal (STR) refer to the factor returns as described on K. French's website. The volatility (VOL) factor is taken from Van Vliet and De Koning (2017) whilst the quality-minus-junk (QMJ) factor is from Asness et al. (2018). Market equity (ME), investment (IA), return on equity (ROE), and expected growth (EG) refer to the factors of Hou et al. (2021). t-statistics are shown in parentheses. α values are reported for the months of January and non-January separately using dummy variables and are expressed in percentage points per month. The sample period is July 31, 1963–December 31, 2021 except for the q-factors, which start on January 31, 1967.

mostly harvesting the size premium.[14]

Model (3) is based on Fama and French's (2015) five-factor (FF5) model including value (HML), profitability (RMW) and investment (CMA) factors in addition to the market and size factors. The coefficient of HML is positive (t-stat of 4.87) but not as powerful as the size factor, indicating that the spread might benefit from a value tilt. RMW has a negative coefficient on the EW–VW spread (t-stat -7.08) whilst CMA is just statistically significant at the 10% level (t-stat -1.83). Notably, the market factor becomes insignificant and negative in this model whilst the 1M-lagged market factor remains significant (t-stat 7.80). The adjusted R^2 of this five-factor model is 79% and hence increases the explained variation by 4 percentage points relative to using the SMB as stand-alone factor.

Models (4) to (5) report alternative factor models: Instead of RMW and CMA they either include the quality-minus-junk (QMJ) factor by Asness, Frazzini, and Pedersen (2019) (in model (5)) or consider the q-factor model by Hou et al. (2021) (model (4)). The EW–VW spread loads negatively on the QMJ factor (t-stat -12.14) highlighting the importance to control for junk amongst the smallest stocks in the EW portfolio. These results are in line with Asness et al. (2018), who analyze the impact of the size effect for different quality factors. Controlling for poor firm quality (or junk) in an EW portfolio is of similar relevance due to its natural exposure to small firms. The q-factor model shows similar loadings to the FF5 factor model, i.e., ME being highly significant (t-stat 34.15) and being complemented by ROE (t-stat -12.69) and EG (t-stat -2.52). Thus, all three models emphasize the impact of size, illiquidity, and quality, but neither model seems to have superior power in explaining the variation of the EW–VW spread with R^2 ranging from 78% for the q-factor model to 81% for the model including QMJ.

Notably, extending the FF5-model with momentum (WML), short-term reversal (STR), and volatility (VOL) factors increases the adjusted R^2 to 85% with WML and VOL having significantly negative coefficients (t-stat -11.05 and -8.66, respectively). The spread loads positively on STR with a t-statistic of 6.48. At the same time, the CMA coefficient turns from negative to positive but is only marginally significant. The size factor SMB as well as

[14]In unreported univariate regressions we confirm the dominant role of SMB as most relevant single factor (adj. R^2 75%), followed by QMJ (34%) and RMW (16%) for the CRSP sample. The S&P 500 sample is driven by SMB (40%), as well as QMJ and WML (both 19%).

the lagged market return (MKT_{t-1}) remain highly significant for all model specifications. These findings resonate with the contrarian rebalancing style of the EW–VW spread which benefits from a short-term reversal effect and momentum underperforming.

With the EW–VW spread return being largely driven by the size factor, it is presumable that other characteristics carry over. One well-researched aspect of the size effect is its seasonality, that is, it is particularly pronounced in January (Keim (1983) and Roll (1983), amongst others). This phenomenon has been linked to investors' year-end tax-loss selling, rebalancing, and cash infusion at the beginning of the year, as well as window dressing by mutual fund managers at the year's end. Indeed, Table 1.1 documents that the EW–VW spread outperformance exclusively accrues in January with a difference in monthly returns of 4.01% vs. -0.05% in non-January months for the CRSP universe. To further rationalize the EW–VW spread, model (7) includes a January dummy (α_{Jan}) alongside MKT, MKT_{t-1}, SMB, HML, RMW, CMA, WML, STR, and VOL. This regression documents strong seasonality in the EW–VW spread return with an average January premium of 193 bps (t-stat 11.85) in the CRSP sample. The other factors remain significant with size still dominating (t-stat 46.34) followed by momentum (t-stat -11.36). The overall adjusted R^2 increases to 87% whilst the baseline α of non-January returns is found insignificant (t-stat 0.16).

Having documented the systematic drivers of the EW–VW spread in the CRSP universe, we next examine whether these results carry over to a more investible universe and thus focus on the S&P 500 index and its constituents. In fact, this selection focuses on the largest and most liquid stocks, with the average S&P 500 stock being almost seven times bigger than the average stock in the CRSP universe (cf. Table 1.1).

First, we focus on the factor model regressions of the S&P 500 EW–VW spread in models (8) to (14) of Table 1.2. Overall, we observe very similar factor sensitivities for the S&P 500 sample compared to the CRSP evidence, yet there are some differences to highlight. On the one hand, the market factor sensitivity MKT for all used factor models is higher than in previous regressions and always statistically significant at the 1% level. At the same time, the lagged market factor becomes insignificant (except for the SIM in model (8)). This is due to the selection of the largest 500 stocks and, hence, a higher exposure to the value-weighted market portfolio and higher liquidity. Note that the relevance of the SMB factor decreases across all models, though it remains highly significant with t-stats ranging from 19.54 to 24.95. Conversely, the value factor HML increases in relevance and ranks third in terms of t-stats (ranging from 7.25 for model (14) to 15.68 in the model including QMJ as quality factor). While the single index model has only slightly lower explanatory power relative to the CRSP case, the average factor model lacks 15 to 35 percentage points in explaining the S&P 500 EW–VW spread compared to the CRSP one. Since all S&P 500 stocks are in the *big* bucket of the size factor SMB by construction, this factor cannot fully address the potential size effects inherent in the S&P 500 universe.

Also, model (14) does not confirm a significant January premium for the S&P 500 universe despite the higher return differences in January, comp. Table 1.1. In fact, the January α is statistically and economically insignificant (t-stat 0.53). The full sample adjusted R^2 stands at 68%. The results are in line with the literature identifying a January premium in small and micro caps but do not confirm such premium for large caps in the S&P 500 universe.

The EW–VW spread over time

To put the observed effects into perspective, we repeat regressions of models (7) and (14) in four subperiods: we specifically look at pre- and post-publication, as well as pre-GFC and post-GFC periods. Table 1.3 depicts the sub-period regression results for the CRSP universe (Panel A) as well as the S&P 500 universe (Panel B). In all subperiods of the CRSP sample, the January α is strictly positive with t-stats ranging from 4.03 (pre-GFC) to 7.95 (post-publication). The subperiod analysis also reveals the decreasing yet highly significant impact of the size factor on the EW–VW spread return over time: Whilst SMB was highly significant in the pre-publication phase (t-stat 44.96) and the model explained almost all return variation (R^2 96%), its impact decreased to its low in the pre-GFC period (t-stat 13.57, R^2 84%). The impact of STR increased over time, from being completely insignificant in the pre-publication period (t-stat 0.56) to becoming a significant driver during the pre-GFC period (t-stat 4.71).

Having screened the spread returns in different subperiods helps to grasp the seasonality in the EW–VW characteristics with the January effect occurring predominantly in the CRSP

Table 1.3: The EW–VW Spread in Different Sub-Periods

$$EW - VW_t = \alpha + \alpha_{Jan} + \beta_1 F_t^1 + ... + \beta_K F_t^K + \epsilon_t \quad for\ K\ factors$$

Period	α	α_{Jan}	MKT	MKT_{t-1}	SMB	HML	RMW	CMA	WML	STR	VOL	Adj.R^2
Panel A: CRSP												
Full sample	0.01	1.93	-0.03	0.09	0.78	0.06	-0.06	0.05	-0.13	0.09	-0.13	0.87
	(0.16)	(11.85)	(-2.52)	(9.50)	(46.34)	(2.95)	(-2.76)	(1.58)	(-11.36)	(5.59)	(-7.94)	
Pre-publication (Expansion)	0.03	1.42	-0.03	0.06	0.90	0.05	-0.08	0.01	-0.05	0.01	-0.08	0.96
	(0.51)	(7.68)	(-1.90)	(5.20)	(44.97)	(1.70)	(-1.97)	(0.25)	(-3.59)	(0.56)	(-3.31)	
Post-publication (Downfall)	-0.18	2.53	-0.07	0.09	0.86	0.08	-0.16	0.20	-0.04	0.07	-0.08	0.87
	(-1.86)	(7.95)	(-2.69)	(4.46)	(22.78)	(1.38)	(-2.55)	(2.98)	(-1.25)	(1.76)	(-1.93)	
Pre-GFC (Recovery)	0.34	2.14	-0.05	0.14	0.65	-0.02	-0.06	0.09	-0.20	0.15	-0.11	0.84
	(2.22)	(4.03)	(-1.09)	(4.74)	(13.57)	(-0.36)	(-0.99)	(1.12)	(-8.21)	(4.71)	(-3.41)	
Post-GFC (Stagnation)	-0.04	1.46	-0.00	0.05	0.73	0.07	-0.15	0.02	-0.08	0.05	-0.20	0.89
	(-0.43)	(5.11)	(-0.07)	(2.71)	(18.82)	(1.72)	(-2.97)	(0.29)	(-2.90)	(1.43)	(-5.31)	
Panel B: SP500												
Full sample	0.03	0.07	0.06	-0.01	0.35	0.13	0.10	0.08	-0.12	0.04	0.03	0.68
	(0.74)	(0.53)	(6.23)	(-1.25)	(24.90)	(7.25)	(5.07)	(2.97)	(-12.65)	(3.09)	(2.02)	
Pre-publication (Expansion)	0.03	-0.68	0.06	-0.01	0.44	0.15	-0.11	0.07	-0.08	0.11	-0.04	0.85
	(0.49)	(-3.44)	(4.11)	(-1.14)	(20.38)	(4.81)	(-2.63)	(1.71)	(-5.09)	(5.15)	(-1.45)	
Post-publication (Downfall)	0.00	-0.34	0.12	-0.03	0.32	0.13	0.08	0.09	-0.15	0.04	-0.02	0.63
	(0.04)	(-1.43)	(5.91)	(-2.39)	(11.30)	(3.07)	(1.65)	(1.73)	(-6.82)	(1.38)	(-0.56)	
Pre-GFC (Recovery)	0.16	0.14	0.09	0.01	0.27	0.10	0.15	0.18	-0.15	-0.00	0.04	0.69
	(1.37)	(0.36)	(2.86)	(0.40)	(7.53)	(2.25)	(3.21)	(2.85)	(-8.25)	(-0.02)	(1.46)	
Post-GFC (Stagnation)	-0.01	0.23	-0.01	-0.01	0.25	0.11	0.02	-0.02	-0.06	0.10	0.05	0.70
	(-0.16)	(1.08)	(-0.78)	(-0.44)	(8.64)	(3.78)	(0.42)	(-0.51)	(-3.11)	(4.06)	(1.65)	

This table reports regression results for the EW–VW spread on the factors MKT, MKT_{t-1}, SMB, HML, RMW, CMA, STR, and VOL, where alphas are estimated for the months of January and non-January separately using dummy variables. Panel A shows results for the CRSP universe; Panel B for the S&P 500 sample. Alphas are presented in percentage points per month. t-stats are in parenthesis. Results are reported over five sample periods: The full sample period (July 1963–December 2021) as well as pre-publication (July 1963–December 1983), post-publication (January 1984–December 1999), pre-GFC (January 2000–December 2009), and post-GFC (January 2010–December 2021) subperiods.

universe. Also, the impact of the systematic factors varies over time with size and negative momentum being the only constant forces whilst the exposure to short-term reversal and profitability comes and goes during different time-periods.[15]

Impact of rebalancing frequency

An important aspect of the EW–VW spread is the need to rebalance frequently to keep the EW component equal-weighted. We next investigate the implications of different rebalancing frequencies in Table 1.4, exploiting rebalancing frequencies ranging from one month (1M, base case) to 60 months (60M). Panel A shows performance characteristics for six differently rebalanced CRSP EW–VW spread portfolios for the full sample period. We observe the highest annualized return for the monthly rebalanced portfolio (3.48%). Interestingly, portfolios rebalanced at the next lower frequency of three months (3M) seem to perform worst in terms of raw as well as risk-adjusted returns (2.15%) whilst performance tends to increase for lower rebalancing frequencies (e.g., 2.94% for 60M). Two-way annualized portfolio turnover decreases monotonically with decreasing rebalancing activities highlighting the impact of drifting weights towards the value-weighted portfolio.

Panel B depicts regression results of the six EW–VW spreads for the full sample period. First, we observe a monotonically declining yet always significant January effect for lower rebalancing frequencies with t-stats ranging from 11.85 (1M) to 2.16 (60M). These results are intuitive as the abnormal January returns are less likely captured if the portfolio is rebalanced at a lower frequency. Next, MKT and MKT_{t-1} indicate clear tendencies of the EW portfolio shifting towards the market portfolio: The negative MKT as well as MKT_{t-1} exposure of the spread lose significance with t-stats shrinking from -2.52 (1M) to -0.24 (60M) and 9.50 (1M) to 4.40 (60M), respectively. Thus, illiquidity and non-synchronous trading concerns become less of an issue given the decreased rebalancing activities.

Another interesting observation is the changing exposure of the EW–VW spread returns

[15] In unreported results we also test EW–VW spread returns within quintile portfolios based on CRSP size breakpoints. They come with similar characteristics as the full sample spreads, albeit the smallest size quintile portfolio shows positive returns in all subperiods with a highly significant January alpha of 218 bps (t-stat 12.11). In terms of factor exposure, the only difference is the negative (albeit mostly insignificant) size exposure for all quintile portfolios but the largest one. This outcome can be related to the missing large cap component of the long-short size factor amongst small and micro-cap stocks, that is, quintiles 1–4.

Table 1.4: Alternative Rebalancing Periods

	Rebalancing frequency					
	1M	3M	6M	12M	36M	60M
Panel A: Performance characteristics						
Ret	3.48	2.15	2.25	2.85	3.00	2.94
Std	11.00	10.44	10.14	9.96	9.69	9.01
Sharpe	0.32	0.21	0.22	0.29	0.31	0.33
MDD	-54.69	-63.61	-63.45	-57.86	-46.21	-39.49
Turnover	1.28	0.72	0.60	0.47	0.27	0.20
Panel B: Regression results						
α	0.01	-0.07	-0.07	-0.06	0.03	0.09
	(0.16)	(-1.56)	(-1.61)	(-1.26)	(0.80)	(2.49)
α_{Jan}	1.93	1.47	1.27	1.14	0.47	0.25
	(11.85)	(9.93)	(8.84)	(7.76)	(3.57)	(2.16)
MKT	-0.03	-0.03	-0.02	-0.01	-0.01	-0.00
	(-2.52)	(-2.67)	(-1.65)	(-0.87)	(-0.68)	(-0.24)
MKT_{t-1}	0.09	0.10	0.09	0.09	0.05	0.03
	(9.50)	(11.29)	(10.71)	(9.76)	(5.77)	(4.40)
SMB	0.78	0.78	0.78	0.78	0.76	0.73
	(46.34)	(50.79)	(52.65)	(51.39)	(56.68)	(61.72)
HML	0.06	0.08	0.09	0.12	0.02	0.00
	(2.95)	(3.96)	(4.76)	(5.90)	(0.95)	(0.08)
RMW	-0.06	-0.07	-0.04	-0.01	-0.15	-0.17
	(-2.76)	(-3.49)	(-1.99)	(-0.33)	(-8.28)	(-10.34)
CMA	0.05	0.06	0.08	0.09	0.10	0.01
	(1.58)	(2.15)	(2.90)	(2.97)	(3.83)	(0.35)
WML	-0.13	-0.10	-0.07	-0.03	0.06	0.05
	(-11.36)	(-9.55)	(-7.40)	(-2.64)	(7.14)	(6.15)
STR	0.09	0.02	-0.01	-0.01	0.01	-0.00
	(5.59)	(1.22)	(-0.57)	(-0.88)	(1.21)	(-0.05)
VOL	-0.13	-0.12	-0.11	-0.11	-0.10	-0.06
	(-7.94)	(-7.79)	(-7.79)	(-7.44)	(-7.88)	(-5.10)
Adj. R^2	0.87	0.88	0.88	0.87	0.89	0.91

This table presents performance characteristics of the CRSP EW–VW spread portfolios with different rebalancing frequencies (Panel A) as well as regression estimates (Panel B) for the full sample. Return, volatility, and 1-month maximum drawdown (MDD) are in percentage terms. Turnover refers to annualized two-way turnover. The EW–VW spread is regressed on the factors MKT, MKT_{t-1}, SMB, HML, RMW, CMA, WML, STR, and VOL, where alphas are estimated for the months of January and non-January separately using dummy variables. Alphas are presented in percentage points per month. t-stats are in parenthesis. The sample period is July 31, 1963–December 31, 2021

to the size factor. The exposure of the size factor decreases with lower rebalancing frequencies (from 0.78 to 0.73) whilst the individual significance of SMB is increased (t-stats ranging from 46.34 to 61.72). This effect can be attributed to the reduced differences of EW and

VW portfolio returns due to drifting weights in the former for lower rebalancing frequencies. At the same time, the statistical fit of the value-weighted size factor SMB is increased the more the EW–VW spread tilts towards a value-weighted portfolio itself.

Moreover, reduced rebalancing frequencies lower and even invert the negative momentum exposure of the EW–VW spread: Whilst monthly rebalancing results in a contrarian strategy with exposure to WML of -0.13 (t-stat -11.36) and positive STR exposure, decreasing the rebalancing frequency to a trend following strategy peaking at a WML exposure of 0.05 (t-stat 6.15) for the 60M portfolio. Overall, the adjusted R^2 is slightly increased with declining rebalancing activities which is potentially linked to the alignment of the EW–VW spread returns with the value-weighted factor construction.

1.4 Investing in the size factor

Given the high correlation between size and the EW–VW spread return as well as size's dominant role in explaining the spread's variation, one would expect the factor's performance to be close to that of the spread returns. In turn, an investor could directly participate in the size premium by simply investing in the EW–VW spread. Unlike SMB, which is difficult to implement, the former can efficiently be implemented using EW market exchange-traded funds (ETFs) which come at low costs compared to rebalancing of a long-short SMB factor portfolio. We analyze the practical implementation of such an EW–VW spread using the S&P 500 as market proxy and compare the capability of this spread to mimic the SMB factor. Additionally, we compare our proposed approach with alternative ways of harvesting the size premium via small cap funds.

In practice, there are different alternatives to long short portfolios based on individual stocks as used for SMB. Starting with the short leg of the portfolio, it is arguably sufficient to short the S&P 500 as such, which market capitalization accounts for the largest proportion of the whole CRSP universe. An investor could therefore choose between selling S&P 500 futures, buying out of the money put options or directly investing in short market ETFs.[16]

[16] All these vehicles come with very specific characteristics and differ in, e.g., liquidity, cost structure and availability for different types of investors. We do not further specify the various vehicles but simply use the same approach to shorting the S&P 500 across the tested alternatives.

In contrast, investing in small stocks is more difficult given their illiquid nature resulting in higher costs. We argue that a cheap way to harvest the size premium is to invest in the EW–VW spread approximated by an EW S&P 500 ETF minus a VW S&P 500 ETF. For benchmarking this suggestion, we focus our analysis of comparable investment options on the iShares Russell 2000 ETF (hereafter R2000 ETF), as well as the DFA US Small Cap and DFA US Micro Cap funds (hereafter referred to as small and micro-cap funds). The former can be generally seen as small cap index whereas the latter two funds primarily invest in small (micro) cap companies whose market capitalizations are generally in the lowest 10% (5%) of total market capitalization. We choose these two funds because the investment process of DFA is heavily influenced by the works of Fama and French. To enable comparing these funds' abilities to harvest the genuine size premium we subtract market returns (as given by the SPX return).

Given the focus on large caps in our approximation, we thus construct another proxy for size effects within the S&P 500 universe (labeled SMBSP) in addition to SMB. This proxy is designed like the original factor by Fama and French (1993) using mid-year median breakpoints to construct the long-short buckets.

Panel A of Table 1.5 depicts the correlation structure of the above return series. We are mostly interested in the ability of the different investment choices to capture the SMB premium. In this regard, the R2000 ETF and DFA small cap fund show a very high correlation of 0.96, and the runner up is DFA's micro-cap fund with a correlation of 0.94. Naturally, these three funds show high cross-correlation in excess of 0.9 as well. In line with earlier factor regressions, the SPW–SPX spread has a lower correlation to SMB but is still reasonably close (0.59).

Of course, we also wish to investigate alignment from a risk-return perspective, see Panel B where we present net performances accounting for annual fees.[17] First, we note that the SMB factor underperforms its S&P 500 counterpart with 151 bps p.a. over the full sample period. This effect is even stronger in the pre-GFC subperiod. The EW–VW spread has

[17]We used annual expenses of 20bps for holding an EW market ETF (e.g. Invesco S&P 500 EW ETF), 19bps for the R2000 ETF, 27bps for the small cap fund and 41bps for the micro-cap fund according to their respective fund prospectus. For shorting the market, we follow the literature (e.g. D'avolio, 2002) and utilize a conservative annual cost of 35 bps.

Table 1.5: Performance Comparison of Size Related Portfolios

Panel A: Correlations

Full sample		SMB	SMBSP	EW –VW	SPW –SPX	R2000 ETF	Small cap	Micro cap
	SMB	1.00						
	SMBSP	0.62	1.00					
	EW–VW	0.79	0.47	1.00				
	SPW–SPX	0.59	0.90	0.58	1.00			
	R2000 ETF	0.96	0.64	0.73	0.63	1.00		
	Small cap	0.96	0.69	0.76	0.68	0.95	1.00	
	Micro cap	0.94	0.57	0.79	0.57	0.91	0.96	1.00

Panel B: Net performance overview

Sample	*Portfolio*	*Ret p.a.*	*Std p.a.*	*Sharpe*	*MaxDD*	*Calmar*	*Sortino*	*CVaR*
FullSample	SMB	3.45	9.49	0.36	−28.59	0.12	0.66	−4.93
	SMBSP	4.96	7.53	0.66	−23.16	0.21	1.13	−4.15
	EW–VW	3.74	10.70	0.35	−35.78	0.10	0.73	−4.82
	SPW–SPX	2.51	5.48	0.46	−21.00	0.12	0.74	−3.13
	R2000 ETF–SPX	1.10	9.71	0.11	−35.27	0.03	0.19	−5.31
	Small cap–SPX	2.34	10.00	0.23	−40.65	0.06	0.40	−5.39
	Micro cap–SPX	2.56	11.24	0.23	−44.64	0.06	0.41	−5.88
Pre-GFC	SMB	7.58	10.05	0.75	−15.54	0.49	1.44	−4.88
	SMBSP	10.29	9.04	1.14	−13.46	0.76	2.20	−4.33
	EW–VW	9.22	12.15	0.76	−27.29	0.34	1.92	−4.53
	SPW–SPX	6.36	6.66	0.96	−13.59	0.47	1.63	−3.57
	R2000 ETF–SPX	4.87	10.32	0.47	−14.80	0.33	0.85	−5.23
	Small cap–SPX	6.95	10.95	0.63	−19.04	0.37	1.14	−5.51
	Micro cap–SPX	7.40	12.34	0.60	−24.42	0.30	1.18	−5.95
Post-GFC	SMB	0.16	8.94	0.02	−28.59	0.01	0.03	−4.85
	SMBSP	0.70	5.81	0.12	−23.16	0.03	0.17	−3.67
	EW–VW	−0.63	9.25	−0.07	−35.78	−0.02	−0.12	−4.90
	SPW–SPX	−0.56	4.13	−0.14	−21.00	−0.03	−0.18	−2.61
	R2000 ETF–SPX	−1.91	9.14	−0.21	−35.27	−0.05	−0.33	−5.20
	Small cap–SPX	−1.34	9.06	−0.15	−40.65	−0.03	−0.24	−5.07
	Micro cap–SPX	−1.31	10.18	−0.13	−44.30	−0.03	−0.21	−5.66

This table shows portfolio correlations (Panel A) and performance characteristics (Panel B) of EW–VW spread and size factor returns for the full CRSP universe (SMB) and the SP500 index (SMBSP), respectively, as well as the three size related portfolios iShares Russell 2000 ETF (R2000 ETF), DFA US Small Cap Portfolio (Small cap) and DFA US Micro Cap Portfolio all minus the SPX return. Return, volatility, maximum drawdown (MaxDD) and expected shortfall (CVaR) are in percentage terms. Results are reported over three sample periods: The full sample period (July 2000–December 2021) and pre-GFC (July 2000–December 2009), and post-GFC (January 2010–December 2021) subperiods.

higher annual returns than its S&P 500 counterpart (3.74% vs. 2.51%), yet coming with the caveat of higher volatility which results in smaller Sharpe ratios (0.35 vs. 0.46) over the full sample. This pattern holds in both subperiods with the exception of negative returns for the SPW–SPX spread in the post-GFC subperiod. The two mutual fund spreads are fairly aligned with that of the R2000 ETF in terms of risk and return (1.10% to 2.56% return p.a. at a volatility of 9.71% to 11.24%). We observe a slight increase in risk-adjusted return the smaller the invested firms become, especially during the recovery phase of the size effect.

However, the three fund spreads show weaker performance characteristics than SMB and the two EW–VW spreads. In fact, they carry more risk at lower returns resulting in smaller risk-adjusted performance (Sharpe ratios around 0.11 to 0.23 vs 0.36 for SMB and 0.35 to 0.46 for the EW–VW spreads). This effect becomes even clearer when looking at the two subperiods: Whilst the size factors and the EW–VW spreads show strong performances during the pre-GFC period with Sharpe ratios ranging from 0.75 (SMB) to 1.14 (SMBSP), all three fund spreads have Sharpe ratios between 0.47 (R2000 ETF) and 0.63 (Small cap). Moreover, during the post-GFC period, size factors and spreads barely exhibit returns different from zero; hence, the resulting Sharpe ratios are close to zero, whilst the three funds even report negative annual returns ranging from -1.31% to -1.91%.

Our performance observations combined with the given correlation structure suggest that the EW-–VW spread is a viable investment alternative for SMB albeit it is not as clean as the decomposition in Section 3 suggested. Nevertheless, the SPW—SPX spread is reasonably close to SMB given its similar performance characteristics during different subperiods as well as its correlation of 0.59. In fact, the EW–VW spread can be considered a cost-efficient alternative to harvest the size factor premium. A closer analysis of subperiods reveals that the characteristics of both EW–VW spreads are close to SMB in terms of risk-adjusted returns whereas the overall correlation of SMB is closer to the Russell 2000 ETF and both DFA funds.

1.5 Conclusion

Historically, the equal-weighted portfolio has outperformed its value-weighted counterpart as well as a variety of other more intricate allocation approaches. In this paper, we identify the key drivers of the EW–VW spread through the lens of different factor models. Focusing on the single index model, we first relate the performance patterns of the VW and EW portfolios to the time-varying market sensitivity β of the average portfolio constituent. However, whilst significant, the single index model cannot explain much of the variation of the EW–VW spread and is thus calling for additional systematic factors in multi-factor regressions of the EW–VW spread.

By design, the EW portfolio is putting more weight into small cap companies which reflects in a massive size exposure relative to a VW portfolio. Also, regular rebalancing to equal weights sees the EW portfolio selling winners and buying losers which is reflected in negative momentum exposures and a positive loading to the Short Term Reversal factor. On average, the EW-VW spread is long higher volatility stocks and thus betting against the Low Volatility anomaly. The over-weighting of small firms also results in negative quality exposure and abnormal high January returns, resonating with the evidence for size-tilted portfolios.

Lastly, we investigate how an investor could participate in the size premium by directly investing in the EW–VW spread. The latter is reasonably close to SMB but comes at lower implementation costs than the long-short factor.

Chapter 2

Compressing the Factor Zoo

This project is joint work with my supervisor Harald Lohre as well as Matthias Hanauer and David Blitz. We thank Amit Goyal, Clint Howard, Pim van Vliet, and participants at the 2023 CEQURA Conference on Advances in Financial and Insurance Risk Management in Munich and the Robeco Research Seminar in Rotterdam for helpful comments and suggestions. This work has been supported by an ESRC NWSSDTP CASE Grant.

2.1 Introduction

To explain the cross-section of stock returns, the asset pricing literature advocates factor models that comprise the factors deemed most representative and relevant. The Capital Asset Pricing Model (CAPM) developed by Sharpe (1964), Lintner (1965), Mossin (1966) and Treynor (1961), is one of the earliest factor models, a one-factor model built around the equity market factor. Despite its theoretical appeal, the CAPM does not perform well in explaining cross-sectional differences in average stock returns. Most prominently, it fails to account for the proper pricing of size (Banz, 1981) and value (Basu, 1977 or Rosenberg, Reid, and Lanstein, 1985) effects, leading Fama and French (1993) to propose a three-factor model consisting of market, size, and value factors. For many years, this model was the industry standard, sometimes augmented with the momentum factor of Jegadeesh and Titman (1993), as in Carhart (1997).

However, over the last 25 years, hundreds of factors have emerged in the literature, all of which allegedly offer a unique new source of return. Cochrane (2011) aptly characterizes the state of play as a 'zoo of factors' that needs to be tamed and structured. As the existing factor models of Fama and French and Carhart (1997) cannot explain many of the new factors, Fama and French (2015, 2018) extend these models to five- and six-factor models by adding investment (Cooper, Gulen, and Schill, 2008) and profitability (Novy-Marx, 2013) factors. These factor models compete with alternative four-factor models such as the Hou, Xue, and Zhang (2015) q-factor model and the Stambaugh and Yuan (2017) mispricing model or the revised six-factor model of Barillas et al. (2020). Although these models use different factors, there seems to be a consensus among leading academics that most of the factor zoo can be explained by parsimonious models consisting of just four to six factors, see Bartram et al., 2021.

Our study of the factor zoo investigates the question of how many factors it takes to compress the factor zoo, i.e., substantially reducing the number of factors without losing (much) information about the tangency portfolio of the entire zoo. To this end, we iteratively identify factors that capture most of the available alpha in the factor zoo. Specifically, the first iteration of our identification strategy augments the CAPM with that factor for which the resulting two-factor model reduces the remaining candidate factor alphas most. The subsequent iteration augments this model further to a three-factor model that captures most of the remaining factor alphas. Sequentially adding factors, we ultimately arrive at a factor model that eliminates all remaining factor alphas. Note that our procedure echoes the approach followed in the early literature, where the CAPM was initially extended by size, value, and momentum factors and later with investment and profitability factors. In contrast, we systematically consider all available candidate factors documented to date until the factor zoo is substantially compressed, leading to alternative paths and insights.

A challenge in analyzing and structuring the factor zoo is to completely reconstruct the existing factors in the literature. To this end, Chen and Zimmermann (2022) as well as Jensen, Kelly, and Pedersen (2023) replicate the vast majority of existing factors and publish open-source databases to facilitate further research. Both studies document that many (if not most) of the proposed factors with high statistical relevance can indeed be replicated, challenging the often-claimed replication crisis in modern finance (Hou, Xue, and Zhang, 2020). Given that Jensen et al. (2023) also provide international factors, our study sources factor data from their database.

Our main findings can be summarized as follows. First, using a comprehensive set of 153 U.S. equity factors, we find that a factor model consisting of 15 factors spans the entire factor zoo. The selected 15 factors originate from 8 out of the 13 factor style clusters, speaking to the heterogeneity of the factor set. Second, iterative factor models also beat common academic models when they contain the same number of factors by selecting alternative value, profitability, investment, or momentum factors or including alternative factor style clusters such as seasonality or short-term reversal. When comparing the existing academic models, we find that the Barillas et al. (2020) revised six-factor model explains most of the available alpha in the factor zoo. Third, when repeating the factor selection over an expanding window and with published factors only, we again recover a diverse set of selected factors. Specifically, newly published factors sometimes supersede older factor definitions, emphasizing the relevance of continuous factor innovation based on new insights or newly available data. Fourth, using equal-weighted factors as opposed to capped value-weighted factors requires more than 30 factors to span the factor zoo, indicating that equal-weighted factors exhibit stronger and more diverse alphas. Finally, applying our factor selection strategy to a set of global factors results in a similar set of selected factors. Although the factor models selected based on global data shrink the alpha for U.S. and World ex U.S. subuniverses, they perform better for the U.S., implying that international factors exhibit larger and more diverse alpha.

We contribute to the literature in several ways. First, we propose a simple yet effective method to identify the important alpha contributors in the factor zoo. The resulting factor sets are relevant from a practitioner's perspective, as they represent the available factor zoo alpha with the minimum number of factors. Our approach differs from previous work on variation in factor returns, e.g., Bessembinder, Burt, and Hrdlicka (2021) and Kozak, Nagel, and Santosh (2018) that mainly investigate the covariance structure in factor returns. These statistical factor studies based on PCA methods typically identify latent factors that describe the covariance structure rather than information about the means (that is, the factors' return level). For instance, consider a hypothetical factor that generates a 1% return every month at zero variance. While this factor would not be considered relevant from a PCA perspective, it is genuinely relevant from a factor premium perspective. In that vein, Lettau and Pelger (2020) develop an alternative 'RP-PCA' approach that incorporates information in the first and second moment of data. Yet, rather than identifying new latent factors in the factor zoo, we are eager to learn about the most relevant factors from an alpha perspective.

Second, we contribute to the debate about the ideal factor model size by consistently identifying 10 to 20 factors over time, depending on the selected statistical significance level. This contrasts with leading academic factor models, which typically only comprise between three and six factors (Barillas et al., 2020; Fama and French, 1993, 2015, 2018; Hou, Xue, and Zhang, 2015; Stambaugh and Yuan, 2017). Interestingly, our results are more in line with the results of studies that apply cross-sectional regressions. For instance, Green, Hand, and Zhang (2017) find that 12 out of 94 characteristics are reliably independent determinants of return among non-microcap stocks, and that 11 of the 12 independent characteristics lie outside prominent benchmark models. Similarly, Jacobs and Müller (2018) find a high degree of dimensionality in international stock returns. Also, the recent evidence from machine learning models indicates that many characteristics matter for predicting individual stock returns (cf., Gu, Kelly, and Xiu, 2020; Hanauer and Kalsbach, 2023; Tobek and Hronec, 2021).

Third, we contribute to the literature on global versus local pricing. Griffin (2002), Fama and French (2012), and Hanauer and Linhart (2015) document that local factors dominate global factors in explaining local return patterns. In contrast, Tobek and Hronec (2021) and Hanauer and Kalsbach (2023) find that the out-of-sample performance of machine learning models for non-U.S. markets is better for global models than for local models. We emphasize the regional impact on factor selection and model construction. While it takes about 6 to 15 factors to span the U.S. factor zoo regardless of the significance level, the global factor zoo is characterized by a similarly sized set of highly significant factors but cannot be compressed to less than 25 to 30 factors at a lower significance level. Lastly, we document a set of global factors that spans the U.S. factors while it needs more factors to span World ex U.S. factors.

The remainder of this paper is structured as follows. Section 2.2 outlines our method for identifying the most important factors in the factor zoo. Section 2.3 presents our empirical results for the U.S. factor zoo, taking into account different weighting schemes and dynamic time periods. Next, we analyze the global factor zoo in Section 2.4 and test the sensitivity of our factor selection method to different regions. Section 2.5 concludes.

2.2 Methodology

2.2.1 Identifying factors that compress the factor zoo

Our goal is to determine the minimum number of factors to explain all factor alphas. From the perspective of a systematic investor, it is worthwhile to identify a factor model that captures as much alpha as possible, since this factor model could guide portfolio allocation for harvesting the underlying factor premiums.

Given the large number of factors put forward in the literature, evaluating competing models is challenging. Valuing the contribution of individual factors vis-à-vis existing alternative factors and quantifying the incremental value added of (potentially) non-nested (i.e., all of the factors in one model are contained in the other model), competing factor models is still an open challenge. Previous work typically differentiates between left-hand-side (LHS) and right-hand-side (RHS) approaches. The former evaluates models by their intercepts (alphas) in time-series regressions of LHS test portfolios' excess returns. Prominent examples for test assets are two-way 5×5 sorts of stocks on size and either book-to-market sorts, momentum, or mispricing (see, e.g., Fama and French, 2015, 2016; Stambaugh and Yuan, 2017) or decile portfolios using various characteristics (see, e.g., Hou, Xue, and Zhang, 2020). However, one limitation of this approach is that the inferences are dependent on the LHS test portfolios and might vary across different test sets (cf., Barillas and Shanken, 2017).

Conversely, Barillas and Shanken (2017) demonstrate that the key in comparing models is how well models price the factors not included in the model and that, surprisingly, the choice of test assets is irrelevant. For nested models, the RHS approach is based on spanning regressions. Specifically, new candidate factors are regressed against the existing model factors to test if they increase the opportunity set. If the corresponding intercept is non-zero, the tested factor contains unexplained information and therefore extends the efficient portfolio frontier. An early proof of the RHS approach is shown, for instance, in Fama (1998), and the approach is applied for model comparisons in Barillas et al. (2020) and Hanauer (2020).

In order to identify a factor model that spans the whole factor zoo from an alpha perspective, we follow a very intuitive and effective nested model approach: We iteratively add new factors to an extending factor model until all remaining alphas in the cross-section of equity factors are rendered insignificant. Our starting point is the CAPM, and we add that factor for which the resulting two-factor model reduces the remaining factor alphas most, measured by the lowest GRS statistic. Please note that this selection criterion is equivalent to selecting the factor with the largest alpha t-stat for the existing model. Once identified, the factor is permanently added to the factor model, and we repeat the procedure based on the resulting augmented factor models until there are no significant contributors left. Formally, the selection strategy can be stated as:

Factor selection steps

Step 1. Set $l := 0$ and start spanning the factor zoo using the CAPM

$$f_i = \alpha_i + \beta_m r_m + \varepsilon_i \qquad i = 1, \ldots, N \tag{2.1}$$

where r_m is the excess market return and N the size of the factor zoo beyond the market.

Step 2. Test $N - l$ different augmented factor models that each add one of the remaining factors, labeled f^{test}, to the model from the previous iteration:

$$f_i = \alpha_i + \beta_m r_m + \sum_{k=1}^{l} \beta_k f_k + \beta^{test} f^{test} + \varepsilon_i \qquad i = 1, \ldots, N - l \tag{2.2}$$

Step 3. Sort the tested factor models based on their explanatory power (as quantified by their GRS statistic, see next section) and select the strongest model.

Step 4. Set $l := l + 1$ and calculate the number of remaining factor alphas $n(\alpha)_{t>x}$ based on the augmented factor model as

$$n(\alpha)_{t>x} = |\{a_i \,|\, t(a_i) > x\}| \qquad i = 1, \ldots, N - l \tag{2.3}$$

where x is the selected significance threshold.

Step 5. Stop if $n(\alpha)_{t>x} = 0$, i.e., if the remaining factors are statistically indifferent from zero. Continue with Step 2 otherwise.

A few things need to be considered when following this iterative nested approach. First, how does one measure the value-add of a tested factor and compare different nested models. Given the linear nature of factor models, it is intuitive to follow a regression-based approach to classify the individual factors' strengths. In the next section, we discuss different metrics used in the literature and rationalize our choice. But note that the above approach also allows for alternative methods to evaluate nested factor models.

Second, a stopping criterion needs to be chosen to effectively pinpoint the number of factors needed to explain all alphas in the factor zoo. We use a straightforward criterion that requires the total number of remaining significant factor alphas to be zero. That is, once a new factor model is identified, we test all remaining factors against this model and determine the alphas for the remaining candidate factors. If the newly added factors are of significance, the number of remaining significant factor alphas should decrease during the process. Alternative criteria could be the significance level of the newly added factor based on the statistical test to identify that factor, i.e., if the new factor does not pass a significance threshold it should not be considered a strong factor and therefore not be added to the model. One caveat of this approach is the large number of regressions needed to identify a factor model that spans the factor zoo. Addressing such data mining concerns and accounting for potential misspecifications, we resort to higher statistical thresholds. Harvey, Liu, and Zhu (2016) deem a t-stat of 3.00 appropriate to account for resulting biases and data mining concerns. Therefore, we run our analysis using the standard thresholds of $t > 1.96$ as well as a more conservative one where $t > 3.00$.

2.2.2 Evaluating factor models

When looking to span the whole factor zoo one is dealing with nested models. A common metric in this field is the GRS statistic of Gibbons, Ross, and Shanken (1989), which produces a test of whether candidate factors help to improve a given model's explanation of expected returns. Specifically, the GRS test investigates whether the alphas of the test assets are jointly different from zero. The GRS test is widely used in empirical finance and has become a standard tool for evaluating the performance of asset pricing models as, e.g., in Fama and French (1996, 2015) and Stambaugh and Yuan (2017). Formally, the empirical GRS statistic is given as follows: Consider an asset pricing model (2.2) with K factors, N test assets, and τ return observations for each time-series. We follow Fama and French (2018) and define the maximum squared Sharpe ratio for the intercepts as

$$Sh^2(\alpha) = \alpha^\intercal \Sigma^{-1} \alpha \tag{2.4}$$

where $\Sigma = e^{\intercal}e/(\tau - K - 1)$ is the covariance matrix of the regression residuals e. The maximum squared Sharpe ratio for the factors of the given model is defined as

$$Sh^2(f) = \overline{f}^{\intercal}\Omega^{-1}\overline{f} \tag{2.5}$$

where $\overline{f}$ is the model's average factor returns and $\Omega = (f-\overline{f})^{\intercal}(f-\overline{f})/(\tau-1)$ is the covariance matrix of the model's factors. The GRS test statistic is calculated as

$$F_{GRS} = \frac{\tau(\tau - N - K)}{N(\tau - K - 1)} \frac{Sh^2(\alpha)}{Sh^2(f)} \tag{2.6}$$

with $F_{GRS} \sim F(N, \tau - N - K)$. The null hypothesis of the GRS test is that all test assets' alphas are strictly equal to zero. If the GRS test statistic is greater than the critical value of the F-distribution at a given significance level, then the null hypothesis is rejected, indicating that the factor model does not adequately explain the variation in test asset returns.

Note that the GRS statistic is crucially determined by the ratio of $Sh^2(\alpha)$ and $Sh^2(f)$. When evaluating factor models, the goal is to identify a model that observes the smallest maximum squared Sharpe ratios for the alphas, and thus captures most of the return variation through its systematic components. Given the relation in equation (2.6), Barillas and Shanken (2017) propose to use the factors' maximum squared Sharpe ratio to evaluate the power of a set of candidate models. Fama and French (2018) further analyze the resulting implications and conclude that the model reducing $Sh^2(\alpha)$ the most is also the model with the highest $Sh^2(f)$, consistent with Barillas and Shanken (2017). Thus, our empirical analysis will not only report GRS statistics and their associated p-values but also $Sh^2(f)$s and average absolute alphas, labeled $Avg|\alpha|$.

2.3 Compressing the factor zoo

2.3.1 Data

Our empirical study is based on the global factor data of Jensen, Kelly, and Pedersen (2023, hereafter JKP), covering 153 factors using data from 93 countries.[1] The set of factors extends the set of factors in Hou, Xue, and Zhang (2020), is similar to that of Chen and Zimmermann (2022), and is thus a meaningful representation of the factor zoo.[2] The JKP database provides one-month holding period factor returns based on the most recent accounting data at a given point in time.

To enable covering all 153 factors in our study, we start our investigation of U.S. capped value-weighted factors in November 1971, and our sample period ends in December 2021. In capped value-weighted factors, stocks are sorted into characteristic terciles each month and the capped value-weighted tercile returns are calculated as the market equity-weighted portfolio returns capped at the NYSE 80th percentile. The factor return is then defined as the high- minus low-tercile return, cf. JKP. This factor construction is designed to create tradable yet balanced portfolios that are neither dominated by mega nor tiny caps. Note that we check for robustness of our main analysis with respect to alternative factor weighting schemes in Section 2.3.5.

Figure 2.1 provides an overview of all factors' annualized alphas based on monthly CAPM regressions. All factors are clustered into 13 categories as identified by JKP based on hierarchical agglomerative clustering (Murtagh and Legendre, 2014), and cluster names are driven by the most representative characteristics. We observe mostly positive annualized alphas for all clusters but the Low Leverage cluster. The average alpha across all factors is 3.51% p.a., and alphas are fairly evenly distributed across and within clusters.

It is not surprising that most of the factors exhibit a significant alpha premium, i.e.,

[1] The factor data is publicly available at https://jkpfactors.com/. An extensive overview of all included factors, their descriptive statistics, as well as the detailed code used to compute them can be found in JKP.

[2] There is a variety of choices that a factor researcher must make in empirical research. For instance, Bessembinder, Burt, and Hrdlicka (2022) highlight the impact of weighting methods and of the number of quantile portfolios underlying the factor portfolios. While they report some differences in the resulting factor portfolio returns across the two mentioned databases, they confirm the statistically significant out-of-sample power of both databases' factors to forecast.

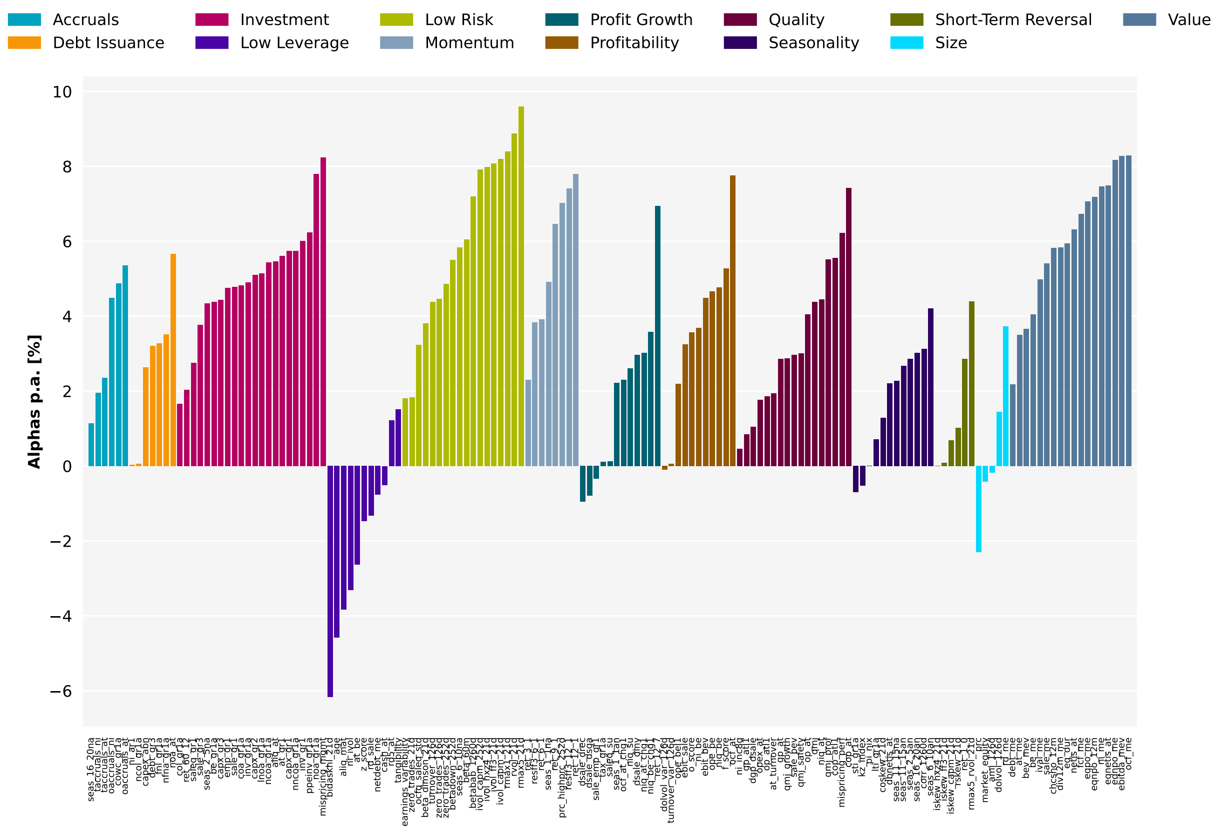


Figure 2.1: Factor Alphas. This figure depicts annualized CAPM-alphas for the U.S. factor zoo. The underlying excess market return is sourced from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. The factor style clusters are based on Jensen, Kelly, and Pedersen (2023). The sample period is November 1971 to December 2021.

indicating incremental power beyond the market return, and thus are deemed to be a relevant factor to begin with. Blitz (2023) further analyzes this set of factors regarding their market risk, performance cyclicality, and inherent seasonal and momentum effects in the cross-section of factor returns. Yet, the question of the incremental alpha contribution of individual factors to the whole factor zoo remains an open question.

2.3.2 Main results

Table 2.1 depicts our main results following the iterative factor selection process described in Section 2.2. We report the selected factors, their associated factor style cluster, the GRS statistics and corresponding p-values, $p(GRS)$, the average absolute intercept $Avg|a|$, the maximum squared Sharpe ratio for the model's factors $Sh^2(f)$, as well as their Sharpe ratios, SR. Columns 10–11 refer to the number of significant factor' alphas after controlling for the specified factor model and thus indicate the incremental explanatory power of the selected factor. The column labels $t > 2$ and $t > 3$ refer to our iterative factor selection with a significance alpha threshold of $t(\alpha) > 1.96$ or $t(\alpha) > 3.00$, respectively. Note that we also report the number of significant factors under common factor models in the subsequent table.

The starting point for our iterative factor selection process is the CAPM model. Based on this one-factor model, we clearly reject the null of the GRS test of all factors' alphas being statistically indistinguishable from zero (GRS statistic of 4.36, p-value 0.00). The CAPM leaves plenty of significant factor alphas, regardless of the selected threshold (105 factors for $t > 2$ and 86 factors for $t > 3$). In the next step, our approach identifies *cash-based operating profits-to-book assets* (cop_at) as the strongest factor in the factor zoo. Adding this quality factor to the market factor still yields a highly significant GRS statistic of 3.54, but the absolute GRS value is clearly reduced. Yet, there are 101 ($t > 2$) or 78 factors ($t > 3$) with an average absolute alpha of 3.94% p.a. in this two-factor model.

The second iteration identifies *change in net operating assets* (noa_gr1a) as the strongest factor amongst the remaining factor zoo contenders. The resulting three-factor model leaves 65 ($t > 2$) or 34 ($t > 3$) significant factor alphas. The average absolute alpha drops to 2.15% p.a. whilst the GRS statistic still remains highly significant (2.98 at a p-value of 0.00).

Iterating further, the factor model increases by construction. Whilst Table 2.1 documents the impact of adding the thirty most relevant factors, it only takes half of that number to span the whole factor zoo. Adding the 15th factor (*highest five days of return scaled by volatility*, $rmax5\_rvol\_21d$) the number of remaining significant alphas drops to zero ($t > 3$). Even with the less strict threshold of $t > 2$, it only takes a total of 18 iterations to render the remaining alphas insignificant. These cut-off numbers are in line with the alternative stopping criterion based on the significance level of the GRS statistic: *highest five days of return scaled by volatility* ($rmax5\_rvol\_21d$) is also the first factor exceeding the 5% significance level with a p-value of 0.09. Overall, these results indicate that it only takes 15 to 18 additional factors to span the factor zoo from an alpha perspective, regardless of the stopping criterion.

Table 2.1: Iterative factor models

No	Factor	Description	Cluster	GRS	p(GRS)	$Avg\|\alpha\|$	$Sh^2(f)$	SR	t>2	t>3
	RMRF	Excess market return	Market	4.36	0.00	3.91	0.02	0.14	105	86
1	cop_at	Cash-based operating profits-to-book assets	Quality	3.54	0.00	3.94	0.15	0.39	101	78
2	noa_gr1a	Change in net operating assets	Investment	2.98	0.00	2.15	0.27	0.51	65	34
3	saleq_gr1	Sales growth (1 quarter)	Investment	2.69	0.00	1.51	0.33	0.58	42	10
4	ival_me	Intrinsic value-to-market	Value	2.49	0.00	1.51	0.39	0.62	39	14
5	resff3_12_1	Residual momentum t-12 to t-1	Momentum	2.31	0.00	1.43	0.44	0.66	35	15
6	seas_6_10an	Years 6-10 lagged returns, annual	Seasonality	2.11	0.00	1.24	0.50	0.71	27	9
7	debt_me	Debt-to-market	Value	1.98	0.00	1.47	0.54	0.74	37	7
8	seas_6_10na	Years 6-10 lagged returns, nonannual	Low Risk	1.87	0.00	1.30	0.58	0.76	25	3
9	zero_trades_252d	Number of zero trades (12M)	Low Risk	1.78	0.00	0.77	0.61	0.78	13	1
10	cowc_gr1a	Change in current operating working capital	Accruals	1.68	0.00	0.88	0.65	0.81	14	3
11	nncoa_gr1a	Change in net noncurrent operating assets	Investment	1.55	0.00	0.70	0.70	0.84	7	1
12	ocf_me	Operating cash flow-to-market	Value	1.48	0.00	0.62	0.73	0.85	5	1
13	zero_trades_21d	Number of zero trades (1M)	Low Risk	1.40	0.01	0.80	0.76	0.87	11	1
14	turnover_126d	Share turnover	Low Risk	1.28	0.03	0.77	0.82	0.90	9	2
15	rmax5_rvol_21d	Highest 5 days of return scaled by volatility	Short-Term Rev.	1.19	0.09	0.63	0.85	0.92	3	0
16	seas_11_15na	Years 11-15 lagged returns, nonannual	Seasonality	1.16	0.14	0.60	0.87	0.93	2	0
17	o_score	Ohlson O-score	Profitability	1.13	0.18	0.67	0.89	0.94	4	0
18	niq_at	Quarterly return on assets	Quality	1.09	0.26	0.59	0.91	0.95	0	0
19	seas_16_20an	Years 16-20 lagged returns, annual	Seasonality	1.07	0.31	0.56	0.92	0.96	0	0
20	ni_ar1	Earnings persistence	Debt Issuance	1.05	0.36	0.56	0.93	0.97	1	0
21	ivol_ff3_21d	Idiosyncratic volatility FF 3-factor model	Low Risk	1.03	0.42	0.48	0.95	0.97	2	0
22	ni_me	Earnings-to-price	Value	0.99	0.52	0.45	0.97	0.98	0	0
23	dsale_dinv	Change sales minus change inventory	Profit Growth	0.97	0.57	0.42	0.98	0.99	0	0
24	ni_be	Return on equity	Profitability	0.96	0.62	0.46	0.99	0.99	1	0
25	noa_at	Net operating assets	Debt Issuance	0.93	0.69	0.46	1.01	1.00	0	0
26	age	Firm age	Low Leverage	0.91	0.73	0.44	1.01	1.01	0	0
27	ret_12_1	Price momentum t-12 to t-1	Momentum	0.90	0.76	0.41	1.02	1.01	0	0
28	aliq_mat	Liquidity of market assets	Low Leverage	0.89	0.78	0.39	1.03	1.02	0	0
29	nfna_gr1a	Change in net financial assets	Debt Issuance	0.88	0.80	0.39	1.04	1.02	0	0
30	at_me	Assets-to-market	Value	0.87	0.83	0.39	1.05	1.02	0	0

This table reports the results for an iterative factor model construction where the k-th iteration augments the model by the factor in row k. It shows the GRS statistic of Gibbons, Ross, and Shanken (1989) and its p-value, p(GRS); the annualised average absolute intercept $Avg|\alpha|$ in percentage, the maximum squared Sharpe ratio for the model's factors $Sh^2(f)$, as well as its Sharpe ratio, SR. Columns 10-11 refer to the number of remaining significant factor alphas after controlling for the specified factor model. $t > 2$ and $t > 3$ control the factor zoo based on the iterative model using a significance alpha threshold of $t(\alpha) > 1.96$ and $t(\alpha) > 3.00$, respectively. The sample period is November 1971 to December 2021.

Against this backdrop, we wonder how our iterative factor selection compares to classic academic factor models. Table 2.2 reports the number of significant factor alphas under well-known models (measured at a significance threshold of $t > 3$). Here, columns FF5 and FF6 refer to the Fama and French (2015) five-factor model, where FF6 augments the latter by a momentum factor. Furthermore, HXZ, BS, and SY refer to the Hou, Xue, and Zhang (2015) q-factor model, the Barillas et al. (2020) revised six-factor model, and the Stambaugh and Yuan (2017) mispricing model, respectively.[3]

Relative to the CAPM (86 alphas with $t > 3.00$), the Fama and French five and six-factor models, the q-factor model, and the mispricing model still leave between 58 and 69 alphas significant. However, the revised six-factor model of Barillas et al. (2020) substantially reduces the number of significant alphas to 33, resonating with it being the only model that includes the *cash-based operating profits-to-book assets* (*cop_at*) factor. The power of the iterative factor model approach is revealed when we compare the academic factor models to the iterative factor models that contain the same number of factors. While the iterative model with four factors merely leaves ten significant factors, the four-factor models of Hou, Xue, and Zhang (2015) and Stambaugh and Yuan (2017) still leave 60 and 55 significant alphas, respectively. Similarly, 14 and 15 remain significant for the iterative model with five and six factors, while for the Fama and French five and six-factor models and the Barillas et al. (2020) revised six-factor model 65, 53, and 29 alphas remain significant, respectively. These results reinforce that the selected factors do not simply represent classic academic factors but are carrying additional information. More specifically, *cash-based operating profits-to-book assets* (*cop_at*) is the only factor from the 15 selected factors that is also contained in one of the common academic factor models, namely the Barillas et al. (2020) revised six-factor model. All the other selected factors represent either alternative value, profitability, investment, or momentum definitions or stem from alternative factor style clusters such as

[3]For consistency, we base these models on those capped value-weighted factors from the JKP database that are most similar to the actual factors used in the original factor model literature. Specifically, the proxies for the SMB, HML, RMW, and CMA factors of Fama and French (2015) are *market equity* (*market_equity*), *book-to-market equity* (*be_me*), *operating profits-to-book equity* (*ope_be*), and *asset growth* (*at_gr1*) from the JKP database. For the models of Hou, Xue, and Zhang (2015) and Barillas et al. (2020) we use *quarterly return on equity*(*niq_be*) and *cash-based operating profits-to-book assets*(*cop_at*) as profitability factors, respectively. The two mispricing factors for the Stambaugh and Yuan (2017) model are *management* (*mispricing_mgmt*) and *performance* (*mispricing_perf*).

Table 2.2: Factor relevance in alternative models

No	Factor	Cluster	$n(\alpha)$							
			t>2	t>3	FF5	FF6	HXZ	BS	SY	C13
	RMRF	Market	105	86	69	58	63	33	58	6
1	cop_at	Quality	101	78	68	57	62	33	57	6
2	noa_gr1a	Investment	65	34	67	56	61	32	56	6
3	saleq_gr1	Investment	42	10	66	55	60	31	55	5
4	ival_me	Value	39	14	65	54	60	30	55	4
5	resff3_12_1	Momentum	35	15	64	53	59	29	54	4
6	seas_6_10an	Seasonality	27	9	63	52	58	28	53	4
7	debt_me	Value	37	7	62	51	57	27	52	4
8	seas_6_10na	Low Risk	25	3	61	51	56	27	52	4
9	zero_trades_252d	Low Risk	13	1	61	51	56	27	52	4
10	cowc_gr1a	Accruals	14	3	60	50	55	26	51	4
11	nncoa_gr1a	Investment	7	1	59	49	54	25	51	4
12	ocf_me	Value	5	1	59	49	54	25	51	4
13	zero_trades_21d	Low Risk	11	1	59	49	54	25	50	4
14	turnover_126d	Low Risk	9	2	59	49	54	25	50	4
15	rmax5_rvol_21d	Short-Term Rev.	3	0	58	49	53	25	50	4
16	seas_11_15na	Seasonality	2	0	57	48	52	24	49	4
17	o_score	Profitability	4	0	57	48	52	24	48	4
18	niq_at	Quality	0	0	56	47	52	24	48	4
19	seas_16_20an	Seasonality	0	0	55	46	51	23	47	4
20	ni_ar1	Debt Issuance	1	0	55	46	51	23	47	4
21	ivol_ff3_21d	Low Risk	2	0	55	46	51	23	47	4
22	ni_me	Value	0	0	55	46	51	23	47	4
23	dsale_dinv	Profit Growth	0	0	54	45	50	22	46	4
24	ni_be	Profitability	1	0	54	45	49	22	45	4
25	noa_at	Debt Issuance	0	0	53	44	48	21	44	4
26	age	Low Leverage	0	0	53	43	47	21	43	4
27	ret_12_1	Momentum	0	0	52	43	46	21	42	4
28	aliq_mat	Low Leverage	0	0	52	42	46	21	42	4
29	nfna_gr1a	Debt Issuance	0	0	51	41	45	20	41	4
30	at_me	Value	0	0	50	40	44	20	41	4

This table reports the results for an iterative factor model construction where the k-th iteration augments the model by the factor in row k. It shows the GRS statistic of Gibbons, Ross, and Shanken (1989) and its p-value, p(GRS); the annualised average absolute intercept $Avg|\alpha|$ in percentage, the maximum squared Sharpe ratio for the model's factors $Sh^2(f)$, as well as its Sharpe ratio, SR. Columns 10-11 refer to the number of remaining significant factor alphas after controlling for the specified factor model. $t > 2$ and $t > 3$ control the factor zoo based on the iterative model using a significance alpha threshold of $t(\alpha) > 1.96$ and $t(\alpha) > 3.00$, respectively. The sample period is November 1971 to December 2021.

seasonality or short-term reversal that offer alpha beyond common factor models (cf., Blitz et al., 2023).

Note that the 15 selected factors emerge from 8 out of the 13 defined factor style categories

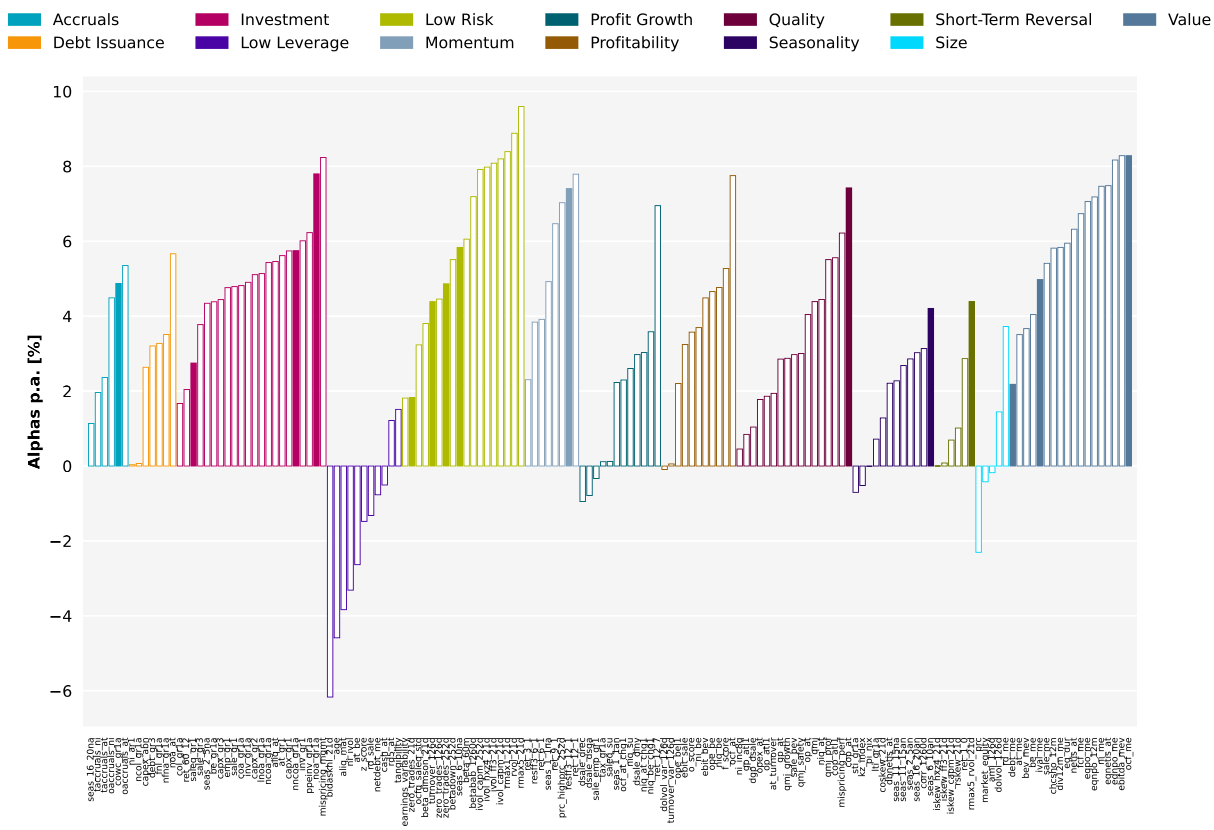


Figure 2.2: Selected Alpha Factors. This figure depicts annualized CAPM-alphas for the U.S. factor zoo. The factors selected by the iterative factor selection process are indicated by full colors; all other factors are whitened. The underlying excess market return is sourced from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. The factor style clusters are based on Jensen, Kelly, and Pedersen (2023). The sample period is November 1971 to December 2021.

and no factor from the remaining five categories is considered, see the highlighted factor bars in Figure 2.2. Moreover, the selected factors are not necessarily those with the highest CAPM alpha in a given factor style cluster; in fact, this only applies to the value, quality, short-term reversal, and seasonality clusters. Notably, whilst five of the eight represented factor clusters merely feature a single factor, the value, low risk, and investment clusters are represented by 3 to 4 factors.

Against this backdrop, we wonder how important it is to go with these selected factors or whether it is sufficient to determine the strongest factor from each of the 13 categories. The last column in Table 2.2 reports results for a 13-factor model consisting of the strongest (largest absolute CAPM alpha) factor per cluster. This model virtually spans the whole zoo, leaving just four factors unexplained. Out of the first 30 factors, it only fails to explain away the alpha from *sales growth for one quarter* ($saleq\_gr1$) and *intrinsic value-to-market* ($ival\_me$), highlighting the power of a cluster spanning model. Nevertheless, the iterative model with 13 additional factors only leaves one alpha significant.

2.3.3 The relevance of factors through time

Having analyzed the full sample evidence in 2.3.2, we wonder about the persistence of the individual factors' relevance through time. One caveat when analyzing the relative strengths of individual factors in the factor zoo is that many factors have only been published along the way and show a weaker post-publication performance (McLean and Pontiff, 2016). To better understand how the set of selected factors evolves through time, we repeat the iterative factor selection but restrict ourselves to the available factors at any point in time. Indeed, many factors that we find to work very well over the whole sample period were not known for many years. For example, the *residual momentum factor* ($resff3\_12\_1$) of Blitz, Huij, and Martens (2011) was only published in 2011 and would, therefore, not have been viable in the first 25 years of the sample. Therefore, we resort to an expanding window analysis using an initial window of 180 months such that we obtain the first out-of-sample observation in December 1986. Each year, we only consider the already published factors when annually running the iterative factor selection and collect the information as presented in Table 2.1. However, note that we stop the iteration at the first occurrence of $n(\alpha)_{t(\alpha)>3} = 0$, i.e., when we arrive at the first factor model that renders all remaining factor alphas insignificant at a threshold of $t > 3.00$.

Figure 2.3 highlights the relevant factors through time, colored by their corresponding factor style cluster. That is, whenever a factor is chosen in the corresponding year's factor model, it is highlighted on the timeline. While the vast majority of factors are either never or rarely included, the top factors from the full sample evidence in Table 2.1 show up prominently, especially over the last 10–15 years.

We observe many factor style clusters to be included in the model for most of the time once a representative factor is published. For instance, the value cluster is present most of the time, but also the momentum cluster is constantly represented since its publication in

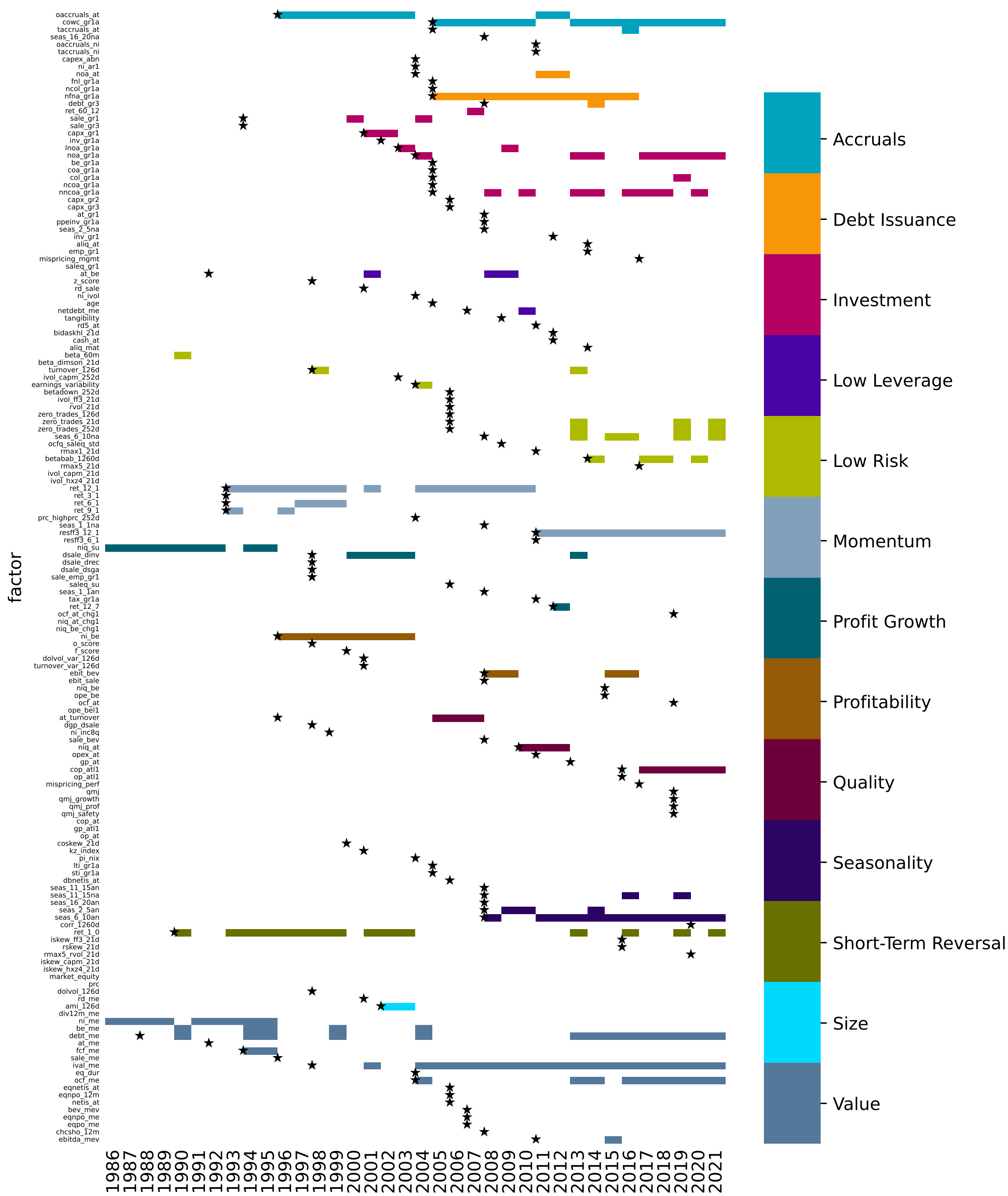


Figure 2.3: Factor Persistence. This figure depicts the expanding development of the iterative model selection process with a significance threshold of $t > 3.00$. Factor are added to the factor zoo on an annual basis once they were published. The publication year for each factor is indicated by a star. The initial window size is 180 months. The sample period is November 30, 1971–December 31, 2021.

Jegadeesh and Titman (1993). Other persistent factor categories are accruals, investment, seasonality, and short-term reversal.

However, many factor style clusters see a change in their representative factors. For example, quality factors have been deemed highly relevant throughout the last three decades, but typically only one quality factor was selected at a time. A similar observation applies to the momentum factor cluster, which is represented by four different factors during the sample period. Notably, the introduction of *residual momentum* ($resff3\_12\_1$) rendered the previously selected classic momentum factor (ret_12_1) insignificant (cf., Blitz, Hanauer, and Vidojevic, 2020). Another example is the accruals cluster. Once published, the new factor *change in current operating working capital* ($cowc\_gr1a$) replaced the *operating accruals* factor ($oaccruals\_at$). These observations emphasize the need and relevance to continuously add and innovate factors and their definitions based on new insights or newly available data.

2.3.4 Rolling window analysis

Given the long-run relevance of different factor style clusters, we next investigate their relevance at shorter time intervals. Therefore, we run a rolling window analysis based on a window size of 180 months that is updated each year. We follow the same iterative factor selection as before and only report the chosen factors that add maximally to a model in leaving the least amount of alphas unexplained.

Figure 2.4 depicts the development of the iterative selection process, aggregated at the factor style cluster level. For each year in the rolling window analysis, we collect the selected factors by style factor cluster. The upper panel of Figure 2.4 plots the number of factors included in each year's model using a cut-off of $t > 2$, whilst the lower panel reports the results for a cut-off of $t > 3$. We report both thresholds to address data-mining concerns while simultaneously gauging the relevance of borderline factors. Also, using a window size of only 180 months naturally raises the bar for factors to exceed a given threshold, relative to the full sample evidence.

The upper panel of Figure 2.4 documents a decrease in the number of selected factors over time. Whilst it took some 15 factors to span the factor zoo in the early years of the sample period, this number decreased to about 8 factors in more recent years. Yet, there are specific factor styles that are persistently relevant through time, including low volatility, seasonality, investment, and quality. While momentum, short-term reversal, and value were included

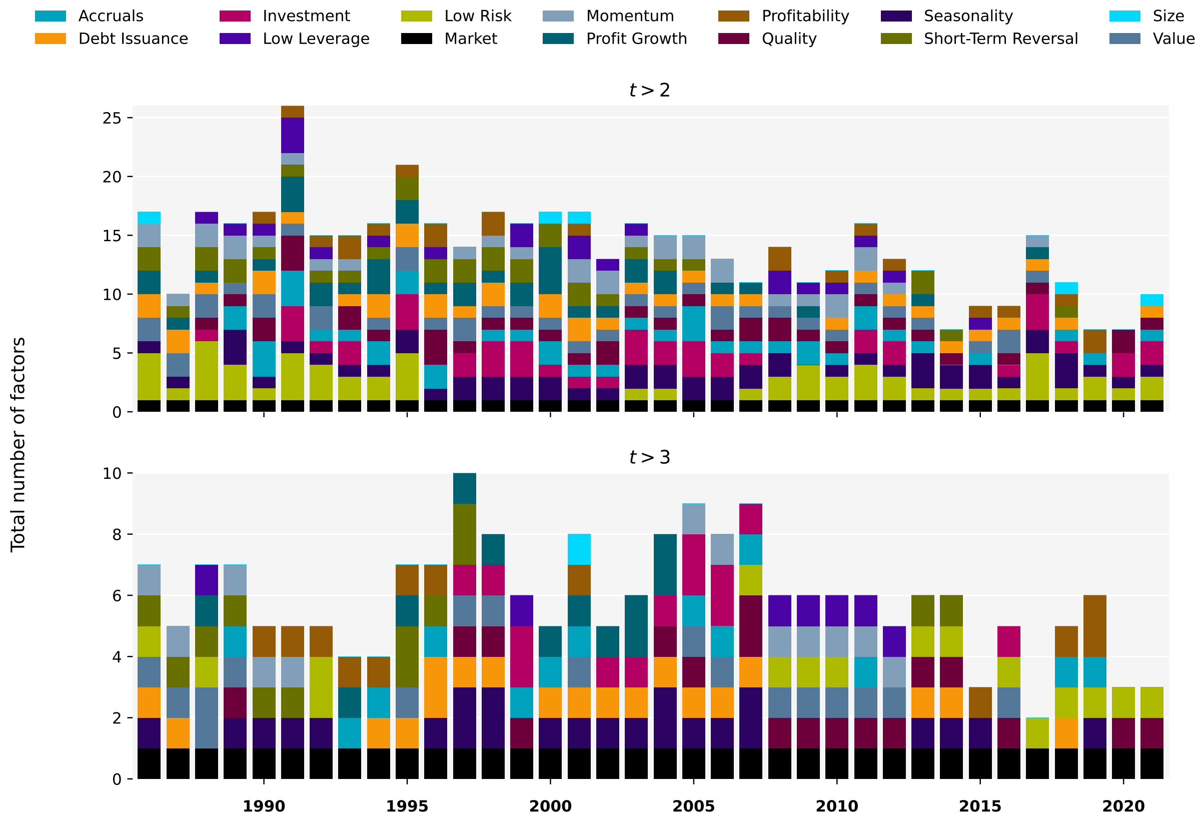


Figure 2.4: Rolling Window Factor Selection. This figure depicts the rolling window outcome of the iterative factor selection using a significance threshold of $t > 1.96$ (upper panel, labelled $t > 2$) and $t > 3.00$ (lower panel, labelled $t > 3$), respectively. The rolling window size is 180 months. The sample period is November 1971 to December 2021.

almost every year until the early 2010s, their relevance has though weakened towards the end of the sample.

Conversely, the lower panel with a threshold of $t > 3$ does only report 4 to 6 factors on average to span the remaining factor zoo. The most relevant factor styles in recent years are quality, low volatility as well as seasonality. Generally, we observe a similar trend in the declining size of the factor models, although the starting models are already quite small compared to models based on the threshold of $t > 2$.

Overall, the represented factor style clusters are slowly changing over time and there is typically some factor representative of the low volatility, seasonality, and quality clusters involved. Interestingly, the classic size factor is rarely chosen and does not seem relevant in spanning other factors' alpha.

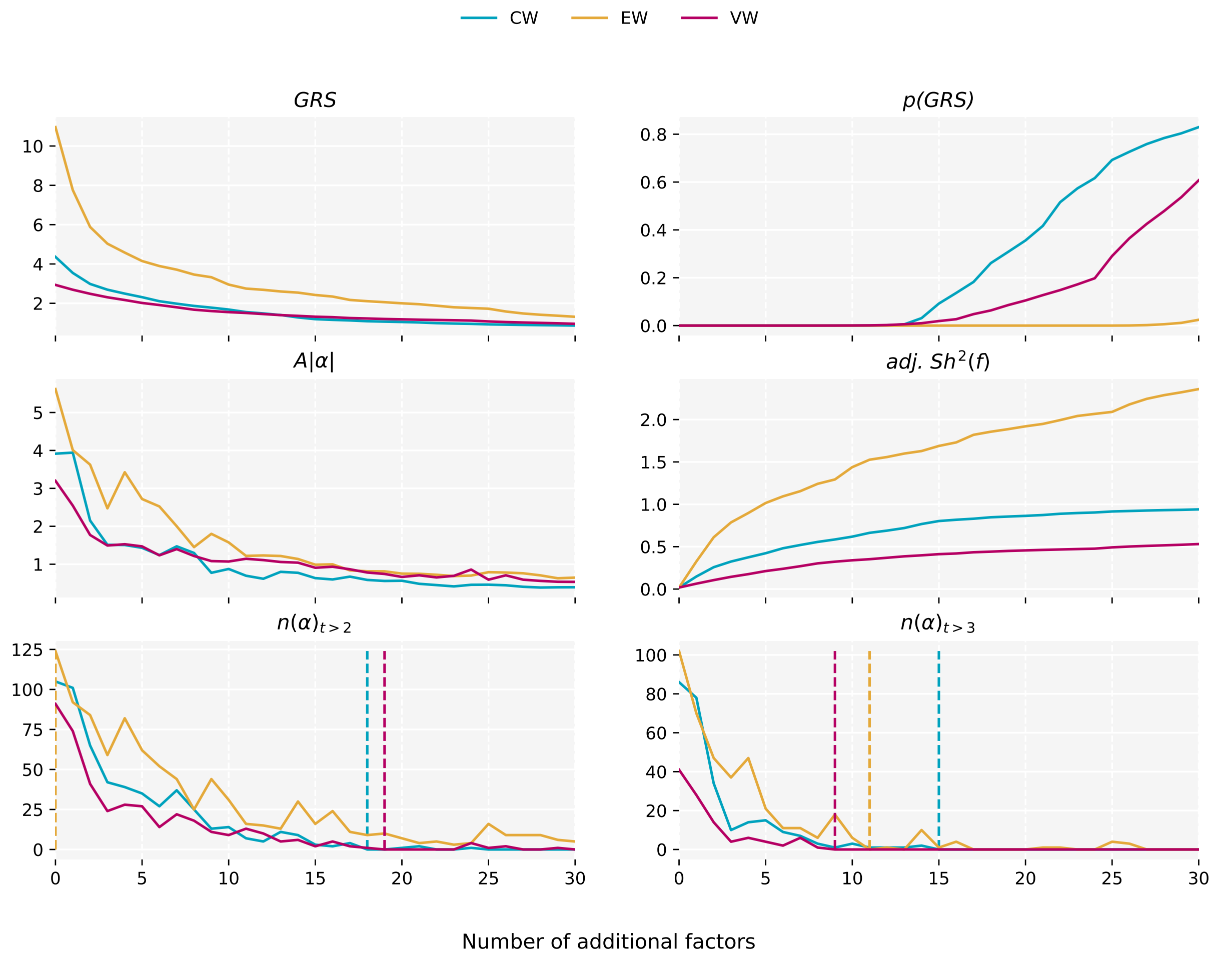


Figure 2.5: Alternative Weighting. This figure depicts the key performance statistics for iterative factor models when based on different weighting schemes. We consider equal-weighting (EW), value-weighting (VW), and capped value-weighting (CW). The number of factors refers to the model building process described in Section 2.2. Average absolute alphas $Avg|\alpha|$ are annualised and in percentage. The vertical dashed lines in the last panel mark the minimum amount of factors needed to explain away the available factor zoo alpha in each weighting setting. The sample period is November 1971 to December 2021.

2.3.5 Robustness regarding alternative weighting schemes

The weighting scheme used to construct factor portfolios can have a big impact, see Bessembinder, Burt, and Hrdlicka (2022) and Soebhag, Van Vliet, and Verwijmeren (2023) amongst others. We thus check for robustness of our results with respect to the three weighting schemes: capped value-weighting (CW), value-weighting (VW), and equal-weighting (EW). We repeat the analysis of Table 2.1 for the different weighting schemes and summarise the corresponding outcome in Figure 2.5.

The first row of Figure 2.5 depicts the development of the GRS statistic and its associated p-value when increasing the number of factors. Whilst the starting GRS statistics for capped value-weighted and value-weighted factor models are low single-digit numbers, the EW factor models come with double-digit numbers. Although the GRS statistic for EW factors quickly declines when increasing the number of factors, p-values suggest significance even at 30 factors, unlike the other two-factor weighting schemes. Comparing CW and VW factor models, we observe CW to pick up in p-values first and cross the 5% threshold when adding the 15th factor, whilst VW takes 18 factors.

Whereas the average absolute alphas, $Avg|\alpha|$, seem to converge towards the same value for all three weighting schemes, the adjusted squared Sharpe ratios, *adj.* $Sh^2(f)$, do not. Indeed, CW and VW factor models each converge towards different *adj.* $Sh^2(f)$, whilst the EW factor models do not yet converge when considering 30 factors. These observations, combined with the results of Fama and French (2018), who generalize that the model with the highest $Sh^2(f)$ must be the best model to minimize the remaining alphas, imply that EW factor models are not confined to a small number of factors to span the entire factor zoo. Put differently, EW factors present a higher and more diverse alpha potential, and thus it takes more factors to span the EW factor zoo.

In the last row of Figure 2.5, we track the number of remaining significant factors for a given factor model. The vertical dashed lines indicate the first occurrence of zero remaining significant factors for a given weighting scheme. Whilst all three weighting schemes reduce the number of remaining significant alphas with increasing size, the EW factor models are sometimes choppy in doing so. That is, increasing the number of factors does not necessarily lead to a decrease in the remaining significant factors in the zoo. Also, it takes 18 and 19 factors for CW and VW factor models to explain away all factor zoo alphas, while EW factor models would take more than 30 factors at a significance threshold of $t > 2$. Notably, this order is almost reversed at a threshold of $t > 3$. Then, VW factor models require 9 additional factors to eliminate significant alphas, followed by 11 and 15 factors for EW and CW, respectively. These results indicate that the selected CW factors are relatively stronger in explaining the CW factor zoo, whereas EW factor models are based on fewer strong factors, and the remainder are more subject to data mining concerns.

2.4 International evidence

2.4.1 Global factor selection

We next broaden our view and investigate whether the U.S. evidence from Table 2.1 carries over to global factors, using international data for 93 different countries. Given the limited availability of some stock-specific measures, we shorten our international sample and focus on the period from August 1993 to December 2021. The sample covers 136 common factors for the three regions World, U.S., and World ex U.S.[4]

Table 2.3 documents the iterative factor selection based on global factor data. Notwithstanding the use of global factors and a shorter sample period, we observe a good overlap in the selected factors compared to the U.S. results in Table 2.1. Out of the first ten selected factors, three are identical ($cop\_at, resff3\_12\_1, cowc\_gr1a$), and the two selected investment factors are close cousins of the U.S. ones. Also, the selected factor style cluster order is very similar (almost identical) for these top factors.

The selection process reveals 11 global factors to span the global factor zoo when enforcing a threshold of $t > 3$. Even at the lower threshold of $t > 2$, it only takes about two dozen factors to span the factor zoo. From a GRS test perspective, it takes between 1 and 2 dozen factors to reject the null of no significant remaining alphas and we can clearly observe the monotonic decline in the GRS statistic for the global factors.

However, the GRS statistics for the U.S. and World ex U.S. samples also decrease almost monotonically, indicating that the global factor models also work for these subsamples. Whilst the models in the U.S. sample have generally lower GRS statistics, resulting in a rejection of the null at about six factors, the models for the rest of the world experience higher GRS values. Note that the shorter sample period of this international analysis already induces a reduction in relevant factors by construction, cf. Figure 2.6 in the next subsection. Although the factor models derived from global data have explanatory power for the World ex U.S. factors, they do a better job on U.S. factors. This observation is probably explained by the much higher alphas that have to be explained as indicated by a GRS statistic of more than 7.03 for international factors compared to a GRS statistic of 2.13 for the U.S.

[4]See Jensen, Kelly, and Pedersen (2023) for a detailed overview of the construction of global factors.

Table 2.3: Global Factor Analysis

No	Factor	Cluster	World GRS	World p(GRS)	World t>2	World t>3	US GRS	US p(GRS)	US t>2	US t>3	World ex US GRS	World ex US p(GRS)	World ex US t>2	World ex US t>3
	market	Market	5.34	0.00	86	50	2.13	0.00	81	44	7.03	0.00	91	56
1	cop_at	Quality	4.11	0.00	57	26	1.77	0.00	71	36	5.43	0.00	73	39
2	ncoa_gr1a	Investment	3.75	0.00	44	14	1.58	0.00	24	7	5.06	0.00	44	15
3	col_gr1a	Investment	3.40	0.00	45	19	1.49	0.00	54	26	4.99	0.00	46	22
4	eq_dur	Value	3.08	0.00	37	16	1.43	0.01	53	19	4.32	0.00	48	17
5	cowc_gr1a	Accruals	2.75	0.00	26	4	1.31	0.04	35	8	4.20	0.00	52	18
6	resff3_12_1	Momentum	2.65	0.00	44	11	1.24	0.08	17	0	3.91	0.00	63	30
7	cash_at	Low Leverage	2.35	0.00	21	3	1.21	0.12	26	8	3.50	0.00	18	5
8	age	Low Leverage	2.19	0.00	18	4	1.21	0.11	27	10	3.46	0.00	32	9
9	dolvol_126d	Size	2.03	0.00	27	4	1.20	0.13	32	12	3.36	0.00	50	19
10	oaccruals_at	Accruals	1.92	0.00	25	4	1.20	0.12	37	19	3.29	0.00	43	18
11	at_be	Low Leverage	1.82	0.00	11	0	1.15	0.18	31	14	3.31	0.00	45	19
12	turnover_var_126d	Profitability	1.75	0.00	8	0	1.16	0.18	32	15	3.05	0.00	42	13
13	nncoa_gr1a	Investment	1.72	0.00	6	0	1.15	0.18	33	14	3.07	0.00	42	13
14	dsale_dinv	Profit Growth	1.67	0.00	6	0	1.16	0.17	34	14	3.09	0.00	43	14
15	iskew_ff3_21d	Short-Term Rev.	1.63	0.00	5	0	1.16	0.17	33	14	3.00	0.00	42	13
16	ret_60_12	Investment	1.59	0.00	3	0	1.13	0.22	33	9	3.02	0.00	42	13
17	rd_sale	Low Leverage	1.54	0.00	3	0	1.13	0.22	36	10	2.61	0.00	20	6
18	mispricing_perf	Quality	1.49	0.01	4	0	1.06	0.35	20	2	2.63	0.00	19	6
19	o_score	Profitability	1.43	0.01	3	0	1.06	0.35	23	2	2.64	0.00	20	5
20	rd5_at	Low Leverage	1.39	0.02	1	0	1.06	0.35	23	2	2.64	0.00	22	5
21	zero_trades_21d	Low Risk	1.34	0.04	9	3	1.06	0.36	17	5	2.57	0.00	19	4
22	zero_trades_126d	Low Risk	1.23	0.10	7	0	0.96	0.58	10	0	2.48	0.00	17	6
23	tangibility	Low Leverage	1.18	0.15	1	0	0.97	0.57	11	0	2.47	0.00	16	6
24	div12m_me	Value	1.13	0.23	1	0	0.96	0.60	5	0	2.44	0.00	17	5
25	be_me	Value	1.07	0.33	1	0	0.96	0.58	6	0	2.22	0.00	8	2
26	prc	Size	1.01	0.47	1	0	0.92	0.68	4	0	2.24	0.00	8	2
27	cop_atl1	Quality	0.98	0.54	0	0	0.91	0.70	6	0	2.21	0.00	6	1
28	coskew_21d	Seasonality	0.96	0.58	0	0	0.90	0.74	6	0	2.23	0.00	6	1
29	qmj_prof	Quality	0.95	0.61	0	0	0.90	0.73	5	0	2.21	0.00	8	1
30	ebit_bev	Profitability	0.93	0.67	0	0	0.88	0.76	4	1	2.23	0.00	8	1

This table reports the results for an iterative factor selection where the k-th iteration augments the model by the factor in row k. The factor selection is based on global factors, and the corresponding factor order is then investigated in the two other regions, U.S. and World ex U.S, using the respective local factors. The table shows the GRS statistic of Gibbons, Ross, and Shanken (1989) and its p-value, p(GRS), as well as the number of remaining significant factor alphas after controlling for the specified factor model. $t > 2$ and $t > 3$ control the factor zoo when based on an iterative model at a significance alpha threshold of $t(\alpha) > 1.96$ and $t(\alpha) > 3$, respectively. The sample period is August 1993 to December 2021 and considers 136 common factors for all three regions.

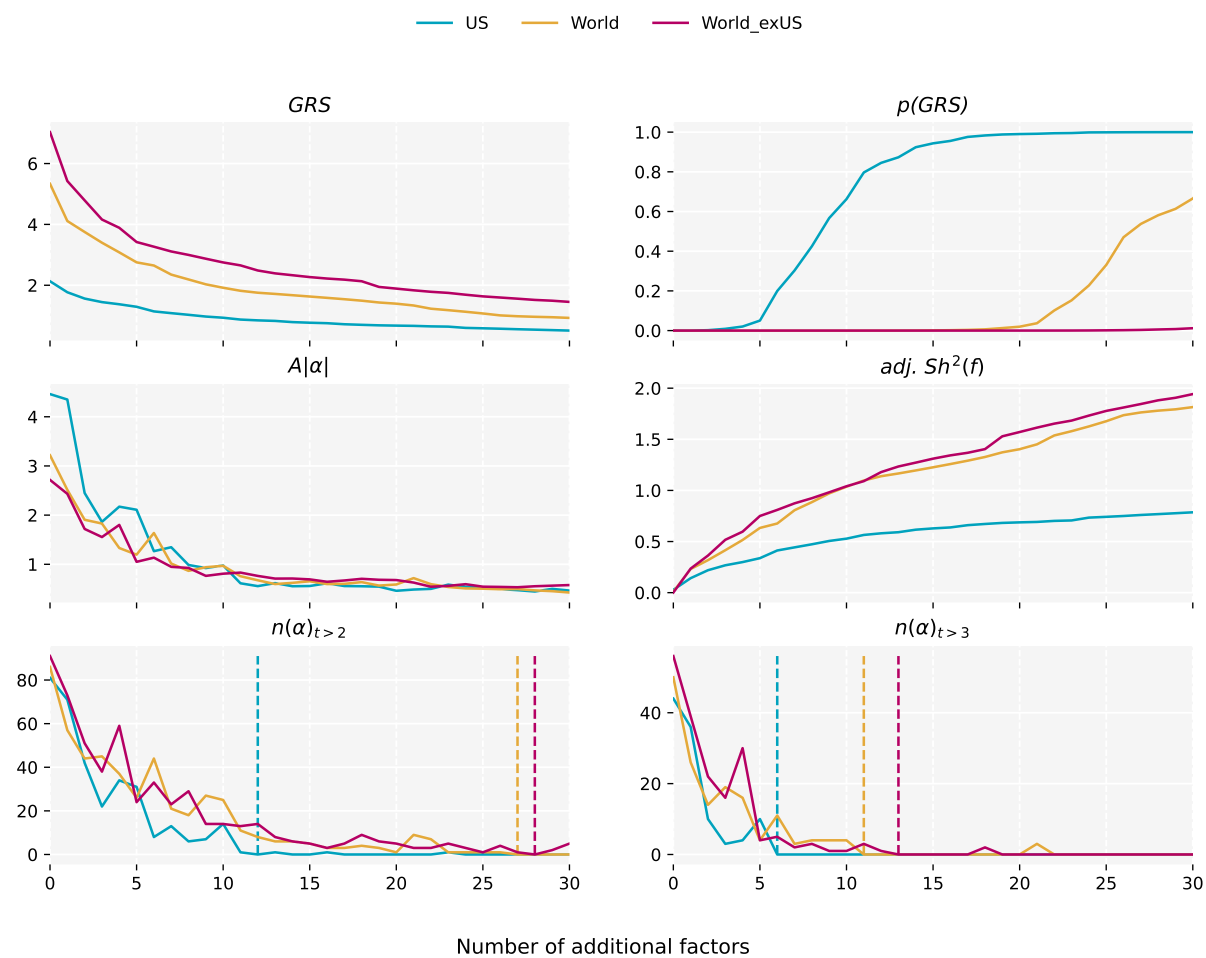


Figure 2.6: Different Regions. This figure depicts the key performance statistics for iterative factor models in different regions. The number and choice of factors refers to the model building process described in Section 2.2. Average absolute alphas $Avg|\alpha|$ are annualised and in percentage points. The vertical dashed lines in the lower panel mark the minimum amount of factors needed to explain the available factor zoo alpha. The sample period is August 1993 to December 2021 and considers 136 common factors for all three regions.

2.4.2 Regional comparisons

So far, global factors have proven relevant for explaining U.S. factor returns, but they lack explanatory power for the World ex US factors. Against the backdrop of Section 2.4.1, we wonder whether local factor models are indeed stronger than global ones. We thus determine the iterative factor model within each region separately and juxtapose the relevant statistics in Figure 2.6 by regions.

The GRS statistics are declining for all three regions by design. However, we clearly see that the World ex-U.S. factors' decline in GRS statistic occurs at a higher level than that of the U.S. and the global factors. As a result, it takes more than 30 factors to reject the null at a significance level of 5% for the World ex U.S. sample whilst U.S. and global models only take 15 and 22 factors, respectively. The differences are primarily driven by the slower convergence of the adjusted squared Sharpe ratios ($adj.\ Sh^2(f)$) in the different regions as shown in the second row of Figure 2.6. Whilst the U.S. factor models seem to converge to some limit in $adj.\ Sh^2(f)$, the other two models' respective lines still have a positive slope at 30 factors. Conversely, the average absolute alphas seem to converge to a statistically insignificant number for all three regions once 30 factors are considered.

Comparing the required model size for spanning all factor zoo alpha we identify another difference between the regions. Whilst the U.S. models have a fairly stable model size with 6 ($t > 3$) to 12 ($t > 2$) highlighting the genuine relevance of the selected factors, the other regions' factors vary in relevance. Out of the 27 (28) factors identified in the global (World ex U.S.) zoo at a threshold of $t > 2$, only 11 (13) are deemed relevant for the higher one at $t > 3$. Thus, these locally selected factor models are even slightly stronger than a universal one, which already helps to span the different regions and especially U.S. factors reasonably well.

2.5 Conclusion

The factor zoo has grown significantly over the last decades as outlined in Cochrane (2011) and Harvey, Liu, and Zhu (2016), highlighting the need to separate sheep from goat factors. To this end, our study investigates the alpha contribution of individual factors in the factor zoo. Specifically, we propose an iterative factor selection strategy to compress the factor zoo, i.e., substantially reducing the number of factors without losing (much) information about the tangency portfolio.

Our results contribute to the literature analyzing the factor zoo and associated factor models. First, we propose a simple yet effective method to identify the relevant alpha contributors in the factor zoo until all remaining alpha sources are exhausted. Our iterative approach operates under different measures such as the GRS test statistic or the maximum squared Sharpe ratio for the factors, $Sh^2(f)$, and is robust across different factor weighting schemes and regions.

Furthermore, we contribute to the debate about the ideal factor model size and identify a persistent size of 10 to 20 factors, depending on the selected statistical significance level. While the academic literature typically advocates parsimonious factor models to explain the cross-section of stock returns, our analysis highlights the need to augment the factor set to fully span the factor zoo from an alpha perspective. Yet, the identified factor models vary over time which is in line with previous literature documenting time variation across factors. Specifically, we identify time-variation in factor representations with new, potentially more robust factors emerging over time. This change in representative factors emphasizes the relevance of factor innovation and diversification of factor metrics to ensure that the selected factors are capable of adapting to the changing factor zoo. However, the identified factor style clusters are quite persistent and emphasize the relevance of diversification across factor styles.

We also emphasize the regional differences in the factor zoo. While it takes about 6 to 15 factors to span the U.S. factor zoo regardless of the significance level, it needs more factors to span World ex U.S. factors, implying that international factors exhibit larger and more diverse alpha. More specifically, we confirm U.S. factor alpha to be well captured by a set of global factors whereas the rest of the world factors seem to be hardly captured. However, local factor models are deemed to explain away most of the factor zoo alpha with less than 15 factors at a significance threshold of $t > 3.00$. Overall, the proposed method allows to span the emerging factor zoo in terms of their alpha contribution across different regions and subperiods, helping investors to rationalize the impact of individual factors in their investment process.

Chapter 3

Macro Factor Investing with Style

This project is joint work with my supervisors Harald Lohre, Mark Shackleton, and Sandra Nolte as well as Scott Hixon and Jay Raol. It is published in the Journal of Portfolio Management (*Journal of Portfolio Management*, 48 (2), 80–104). We thank Tarun Gupta, Scott Wolle, and participants at the 2021 Frontiers of Factor Investing Conference in Lancaster, the 2021 CEQURA Conference on Advances in Financial and Insurance Risk Management in Munich, and the 2020 Global Research Meeting of Invesco Quantitative Strategies in New York for helpful comments and suggestions. This work has been supported by an ESRC NWSSDTP CASE Grant.

3.1 Introduction

Diversification is a central tenet in modern asset management, dating back to the seminal work of Markowitz (1952). Consequently, many investors seek to diversify their portfolios by asset class, despite the fact they can be highly correlated during bear markets and especially during extreme market shocks. The phenomenon of asymmetric spikes in correlation for extreme downside moves rather than upside swings is well recognized (e.g., Ang and Chen, 2002; Longin and Solnik, 2001) and poses a challenge to the effectiveness of diversification. If traditional diversification across asset classes fails when it is needed the most, then identifying the latent drivers behind such comovements can help investors better navigate these periods. We propose using macroeconomic factors, which are natural candidates to describe economic scenarios and explain return variation across asset classes and style factors.

Specifically, we investigate macro factor strategies that aim for high diversification across a parsimonious set of macroeconomic factors. In modern asset pricing theory, macroeconomic factors represent different market states. Investing in assets with high exposure to these market states is compensated for with associated risk premia. Macro factor investing involves several key design choices: First, we need to identify the factors. Some researchers rely on macroeconomic state variables, while others use statistical factors from a principal component analysis of the given investment universe. Second, it is not usually straightforward to invest in macroeconomic factors, because they follow economic indicators that lack corresponding investment vehicles. If direct investment is cumbersome, we may need to construct macro factor-mimicking portfolios (MFMPs) that render the targeted macro factor investable by mapping it on investable assets. In this paper, we focus on the three salient macroeconomic factors, *Growth*, *Inflation*, and *Defensive*, that enable us to navigate a multitude of economic scenarios relevant to large groups of investors.

Our work is related to Amato and Lohre (2020), who compare the explanatory power of pure macroeconomic state variables to that of pure risk factors in building diversified macro factor allocations. While the authors derive macro factor-mimicking portfolios based on a set of traditional asset classes, we extend the macro factor lens to encompass the salient style factors in the investigated asset classes. Indeed, we build more diversified macro factor-mimicking portfolios by isolating macroeconomic exposure and investigating their merits in corresponding macro factor completion analyses. One necessary ingredient for macro factor completion is an optimal macro factor allocation toward which one can tilt a given portfolio allocation. To this end, we build specifically on diversified risk parity (DRP) strategies, following Lohre, Opfer, and Ország (2014) and Dichtl et al. (2021).

We contribute to the literature in several ways. First, we build out diversified MFMPs based on asset classes and style factors. Including style factors enhances the ability of the mimicking portfolios to capture not only asset-specific but also style factor-related responses to macroeconomic shifts. Second, we investigate the integrated management of style and market factors through a macro factor lens. Controlling portfolio risk exposure helps reduce tail risks and mitigates extreme downfalls during economic shocks. Third, we tailor macro factor completion strategies according to individual investor objectives.

Our results confirm the importance of controlling for individual asset and macroeconomic risk. Balanced exposure to three macroeconomic risk sources enhances the portfolios' risk-return profiles given the associated reduction in maximum drawdowns. Macro factor completion enables a significant increase in effective bets, and, hence, diversification. For example, the completion strategies of a 60/40 stock-bond portfolio lead to an increase in the average Sharpe ratio from 0.54 to between 0.59 and 1.01 depending on the nature of the overlay.

The remainder of this paper is structured as follows. Section 3.2 discusses the rationale for using a multi-asset multi-factor setup, reviews the relevant macroeconomic factor literature, and explains our choice of macro factors. Section 3.3 describes the theoretical framework, recapping the orthogonalization technique introduced by Meucci, Santangelo, and Deguest (2015), and deriving the corresponding factor risk parity allocations. Section 3.4 presents our empirical results, including the analysis of macro factor sensitivities of individual assets, as well as the portfolio implications of the macro factor completion framework. It also discusses our robustness tests and further applications from broadening the set of macro factors. Section 3.6 concludes.

3.2 The nature of macro factors

3.2.1 Factor models and macro factor allocation

Since the introduction of the capital asset pricing model (CAPM),[1] a variety of factor models have been developed to represent an asset's expected return as a linear combination of systematic risk factors. Ross's (1976) arbitrage pricing theory (APT) is a seminal contribution. Under APT, the returns $\boldsymbol{R} \in \mathbb{R}^{N \times 1}$ of N risky assets follow a factor intensity structure expressed as:

$$\boldsymbol{R} = \boldsymbol{B} \cdot \boldsymbol{F} + \boldsymbol{\varepsilon}, \tag{3.1}$$

where $\boldsymbol{F} \in \mathbb{R}^{K \times 1}$ represents the returns of K systematic factors with respective factor loadings $\boldsymbol{B} \in \mathbb{R}^{N \times K}$ and asset-specific idiosyncratic risks $\boldsymbol{\varepsilon} \in \mathbb{R}^{N \times 1}$, which are assumed to be uncorrelated across assets and factors and have zero mean. The expected asset returns can be expressed in terms of factor sensitivities, so that:

$$\mathbb{E}\left(\boldsymbol{R}\right) = \boldsymbol{r_f} + \boldsymbol{B} \cdot \boldsymbol{RP}, \tag{3.2}$$

with $\boldsymbol{r_f} \in \mathbb{R}^{N \times 1}$ denoting the risk-free rate, and $\boldsymbol{RP} \in \mathbb{R}^{K \times 1}$ denoting the risk premia associated with the corresponding systematic factors. Several factor models follow this paradigm, e.g., Fama and French (1993), Carhart (1997), Hou, Xue, and Zhang (2015), and Fama and French (2015).

However, many investment processes focus on the optimal allocation of asset classes, instead of the underlying factors and asset allocators that would typically decompose portfolio returns into systematic risk drivers and idiosyncratic risk ex post. Even if asset-specific factors (hereafter referred to as style factors) are considered during the allocation process, it is not usually systematic across multiple asset classes and style factors. A macro factor framework can help provide such rigor.

One method for defining macroeconomic factors aims to capture different states of the economy based on fundamental data such as output or employment rates (e.g., Chen, Roll, and Ross, 1986). This approach is transparent and easy to interpret, but comes with a

[1] See Sharpe (1964), Lintner (1965), Mossin (1966) and Treynor (1961).

caveat: Many key economic figures are generated through economic surveys, which are at best proxies for the real economic state, and are observed only at low frequencies. Further, post-publication revisions and/or interpolation of low-frequency data points make it difficult to rely on such factors for related investment processes. As a result, such macroeconomic state factors are better used for long-term decision making and economic forecasting than for short-term portfolio allocation decisions.

Another method uses a statistical approach. Asset returns are decomposed via techniques that extract principal components, which explain most of the assets' return variation (see, e.g., Bass, Gladstone, and Ang, 2017). Generally, statistical factors tend to have higher explanatory power but can lack economic interpretation due to their mechanical construction. However, regardless of the definition of macroeconomic factors, it is important to construct investable macro factors (referred to as macro factor-mimicking portfolios, or MFMPs). In other words, MFMPs are portfolios of investable assets that aim to replicate the original factors' behavior.

There are different ways to construct these portfolios, from two-pass cross-sectional regression models (e.g., Fama and MacBeth, 1973) and maximal correlation applications (e.g., Huberman, Kandel, and Stambaugh, 1987; Lamont, 2001), to machine learning approaches (Jurczenko and Teiletche, 2022). Note that factor-mimicking portfolios may come with high turnover and transaction costs, which derive from estimation errors in the respective factor loadings (rather than actual allocation changes). Therefore, deriving stable MFMPs is crucial for investing efficiently in macro factors.

In this vein, we present an intuitive approach to managing macro risk exposure, which lends itself naturally to the practice of macro factor portfolio management.

3.2.2 Multi-asset multi-factor investing and macro factors

Investment universe

To illustrate our approach, we work with a clearly defined number of assets, style factors, and macro factors. Thus, we can explicitly address and manage the implicit macro factor

Table 3.1: Data Description

Name	Description	Ticker	Source	Construction details
Equities				
ACWI	MSCI ACWI Net TR Local index	NDLEACWF	Bloomberg	
US-ACWI	MSCI USA TR USD Index minus MSCI ACWI Net TR Local Index	NDDLUS, NDLEACWF	Bloomberg	NDDLUS - NDLEACWF
EAFE-ACWI	MSCI EAFE TR LCL Index minus MSCI ACWI Net TR Local Index	NDDLEAFE, NDLEACWF	Bloomberg	NDDLEAFE - NDLEACWF
EM-ACWI	MSCI EM TR LCL Index minus MSCI ACWI Net TR Local Index	NDLEEGF, NDLEACWF	Bloomberg	NDLEEGF - NDLEACWF
Cyclicals-Defensives	ACWI CYCLICAL SECTORS- ACWI DEFENSIVE SECTORS	MXCXDRN	Bloomberg	
Equity LowVol	MSCI ACWI Minimum Volatility USD minus MSCI ACWI Net TR USD Index	M00IWD$O, NDUEACWF	Bloomberg	M00IWD$O - vol.adj. NDUEACWF
Equity Quality	MSCI ACWI Quality USD minus MSCI ACWI Net TR USD Index	M1WDQU, NDUEACWF	Bloomberg	M1WDQU - vol.adj. NDUEACWF
Equity Momentum	MSCI ACWI Momentum USD minus MSCI ACWI Net TR USD Index	M1WD000$, NDUEACWF	Bloomberg	M1WD000$ - vol.adj. NDUEACWF
Equity Value	MSCI ACWI Value USD minus MSCI ACWI Net TR USD Index	M1WD000V, NDUEACWF	Bloomberg	M1WD000V - vol.adj. NDUEACWF
Fixed Income				
US 10Y Tsy	Bloomberg Barclays US Treasury Bellwethers 10 Year TR Index Unhedged USD	BW10TRUU	Bloomberg	
Cash	USD 3 Month T-Bill	USGG3M	Bloomberg	
TIPS	US TIPS TR	I01551US	Bloomberg	
IG Credit	Bloomberg Barclays US Agg Corp excess return	LUACER	Bloomberg	
HY Credit	Bloomberg Barclays US Corporat HY excess return	LF98ER	Bloomberg	
EM Credit	J.P. Morgan EMBI Global TR minus US Treasury	JPEIGLBL, LUATTRUU	Bloomberg	JPEIGLBL - vol.adj. LUATTRUU
Rates Carry	Goldman Sachs Rates Carry Strategy	GSIRCA03	GS	
Rates Quality	Goldman Sachs Rates Quality Strategy	GS Interest Rates Curve C0210	GS	
Rates Momentum	Goldman Sachs Rates Momentum Strategy	GSIRTR03	GS	
Rates Value	Goldman Sachs Rates Value Strategy	GSIRVA03	GS	
Commodities				
Precious Metals	Bloomberg Precious Metals Subindex	BCOMPR	Bloomberg	
Industrial Metals	Bloomberg Industrial Metals Subindex	BCOMIN	Bloomberg	
Energy	Bloomberg Energy Subindex	BCOMEN	Bloomberg	
Agriculture	Bloomberg Agriculture Subindex	BCOMAG	Bloomberg	
Cmdty Carry	Goldman Sachs Commodity Carry Strategy	GS Macro Carry Index RP14	GS	
Cmdty Quality	Bloomberg Roll Select Commodity Index minus Bloomberg Commodity Index	BCOMRST, BCOMTR	Bloomberg	BCOMRST - BCOMTR
Cmdty Momentum	Goldman Sachs Commodity Momentum Strategy	GS Macro Momentum Index RP15	GS	
Cmdty Value	Goldman Sachs Commodity Value Strategy	GS Commodity COT Strategy COT3	GS	
Currencies				
Developed Markets	MSCI EAFE Currency USD Index	MXEA0CX0	Bloomberg	
Emerging Markets	MSCI Emerging Markets Currency USD Index	MXEF0CX0	Bloomberg	
FX Carry	Goldman Sachs FX Carry Strategy	GS FX Carry C0115	GS	
FX Momentum	Goldman Sachs FX Momentum Strategy	GS FX Trend C0038	GS	
FX Value	Goldman Sachs FX Value Strategy	GS FX Value C0114	GS	
Macro factors				
Growth	MSCI ACWI Net TR Local index	NDLEACWF	Bloomberg	
Defensive	Bloomberg Barclays US Treasury Bellwethers 10 Year TR Index Unhedged USD	BW10TRUU	Bloomberg	
Inflation	USTreasuryTIP minus BBUSTreasury	SPBDUP3T, LT01TRUU	Bloomberg	SPBDUP3T - LT01TRUU
Commodity	Bloomberg Commodity Index	BCOM	Bloomberg	
EM	MSCI Emerging Market Index minus MSCI World Index	MXEF, NDDUWI	Bloomberg	MXEF - NDDUWI
Credit	Bloomberg Barclays US Agg Corp plus Bloomberg Barclays US Corporat HY	LUACER, LF98ER	Bloomberg	0.8*LUACER + 0.2*LF98ER
FX	MSCI EAFE CCY USD Index plus MSCI Emerging Markets CCY USD Index	MXEA0CX0, MXEF0CX0	Bloomberg	0.8*MXEA0CX0 + 0.2*MXEF0CX0

This table describes the construction of all asset class and style factors, as well as the underlying indices and identifiers from Bloomberg and Goldman Sachs (GS). Note that equity style factors and EM Credit are neutralized based on volatility-adjusted indices that are rescaled using an EWMA volatility forecast for the two underlying indices. The EWMA forecast covers at least twelve months of data.

sensitivities that come with investing across salient drivers of asset returns. We focus on a broad multi-asset universe that includes *Equities*, *Fixed Income*, *Commodities*, and *Currencies*. Our sample period is from January 2001 to June 2021, and we use monthly data from Bloomberg and Goldman Sachs. A detailed explanation of assets and style factors, along with the specific tickers used, is in Table 3.1. All indices are measured in local currency returns, and we report total returns of multi-asset multi-factor strategies from a U.S. investor perspective.

Table 3.2 gives an overview of the descriptive statistics for all asset class and style factor returns that we report in excess of three-month U.S. T-bills. For Equities, we consider five assets as well as four style factors. The MSCI All Country World Index (ACWI) represents global equity, with an annualized excess return of 5.69% and volatility of 14.27%. US-ACWI, EAFE-ACWI, and EM-ACWI are the active returns of the corresponding regions relative to the ACWI. Each is measured as the region's return minus the MSCI ACWI return. Their active returns range from -2.10% p.a. for EAFE to 4.03% p.a. for Emerging Markets, with annualized volatilities of between 3.45% for the U.S. and 8.95% for Emerging Markets. Cyclical-Defensive is the spread of U.S. cyclical equities over defensive. Regarding the four equity style factors, Quality, Momentum, Value, and Low Volatility, we compute the spread between the corresponding MSCI ACWI style factor index and the MSCI ACWI. The Quality, Momentum, and Low Volatility factors exhibit annual excess returns of 3.11%, 4.73%, and 2.48%, respectively. Conversely, Value underperformed on average by -1.72% p.a. Note that equity factors come at reduced volatility when compared to asset class volatility: 3.23% for Value vs. 7.46% for Momentum.

For Fixed Income, we consider U.S. 10Y Treasuries, TIPS, as well as IG, HY, and EM credit. The former two can be considered safe haven assets, with annualized excess returns of 3.41% and 4.00%, respectively. The three credit spreads are computed as excess returns over Treasuries. While IG and HY Credit come at a premium of 1.18% and 3.76% p.a., EM Credit is slightly positive at 0.15%.

We also consider the four rates style factors: Carry, Quality, Momentum, and Value. Only two have positive excess returns: 1.56% for Rates Carry, and 2.28% for Rates Momentum. Value and Quality have negative excess returns, with -0.53% and -0.14%, respectively.

Table 3.2: Descriptive Statistic of Assets, Style, and Macro Factors

	Ret p.a.	Vol p.a.	SR	t-stat	Min	Max	MaxDD
Equities							
ACWI	5.69	14.27	0.40	1.80	-16.98	11.42	-51.86
US-ACWI	1.28	3.45	0.37	1.68	-2.80	2.81	-16.64
EAFE-ACWI	-2.10	4.31	-0.49	-2.21	-5.76	3.35	-38.51
EM-ACWI	4.03	8.95	0.45	2.03	-6.13	8.05	-35.81
Cyclical-Defensive	0.56	11.57	0.05	0.22	-12.2	13.63	-57.31
Equity LowVol	2.48	4.76	0.52	2.35	-4.34	4.89	-15.70
Equity Quality	3.11	3.48	0.89	4.04	-2.20	4.81	-10.78
Equity Momentum	4.73	7.46	0.63	2.86	-7.71	12.91	-17.82
Equity Value	-1.72	3.23	-0.53	-2.40	-3.44	2.49	-35.30
Fixed Income							
US10YTsy	3.41	7.25	0.47	2.13	-7.15	9.03	-12.42
TIPS	4.00	5.70	0.70	3.17	-8.73	5.83	-13.03
IG Credit	1.18	5.27	0.22	1.01	-10.40	5.39	-24.2
HY Credit	3.76	10.64	0.35	1.60	-16.50	13.21	-44.76
EM Credit	0.15	12.44	0.01	0.05	-17.49	9.42	-42.71
Rates Carry	1.56	3.79	0.41	1.86	-4.13	2.98	-12.21
Rates Quality	-0.14	2.94	-0.05	-0.22	-2.34	3.82	-19.87
Rates Momentum	2.28	4.37	0.52	2.35	-3.32	3.71	-10.68
Rates Value	-0.53	3.80	-0.14	-0.63	-3.68	2.73	-21.84
Commodities							
Precious Metals	7.51	19.30	0.39	1.76	-19.01	14.45	-51.18
Industrial Metals	4.57	21.63	0.21	0.95	-26.97	21.35	-67.00
Energy	-6.21	29.36	-0.21	-0.96	-35.13	24.78	-97.25
Agriculture	0.09	19.94	0.00	0.02	-18.97	15.85	-67.60
Cmdty Carry	2.74	8.14	0.34	1.52	-6.93	7.40	-35.25
Cmdty Quality	3.92	2.70	1.45	6.56	-2.49	4.74	-3.74
Cmdty Momentum	-1.06	9.67	-0.11	-0.49	-7.94	9.38	-50.54
Cmdty Value	3.29	7.50	0.44	1.98	-5.57	6.84	-14.75
Currencies							
DM Basket	1.00	7.02	0.14	0.64	-5.49	6.28	-28.44
EM Basket	2.70	6.36	0.42	1.91	-7.56	5.17	-19.21
FX Carry	3.34	6.27	0.53	2.41	-8.15	7.11	-16.59
FX Momentum	1.70	4.87	0.35	1.57	-3.27	5.24	-12.20
FX Value	1.07	4.50	0.24	1.08	-6.43	4.77	-18.82

To be continued on the next page

Table 3.2 – (continued)

	Ret p.a.	Vol p.a.	SR	t-stat	Min	Max	MaxDD
Macro factors							
Growth	5.69	14.27	0.40	1.80	-16.98	11.42	-51.86
Defensive	3.41	7.25	0.47	2.13	-7.15	9.03	-12.42
Inflation	1.21	2.65	0.46	2.07	-6.16	2.49	-10.87
Commodity	-0.83	15.89	-0.05	-0.24	-21.38	12.98	-75.52
FX	1.34	6.56	0.20	0.92	-5.81	5.90	-25.67
EM	1.15	11.49	0.10	0.45	-9.05	7.93	-60.66
Credit	1.69	6.13	0.28	1.25	-10.98	6.95	-28.66

The table shows descriptive statistics of excess returns for asset classes, style factors, and macro factors. Min and Max denote the lowest and highest monthly excess return during the sample period. SR is the corresponding Sharpe ratio and t-stat reports the t-statistic for testing the the null hypothesis that the SR equals 0. Return, volatility, Min, Max, and Maximum Drawdown (MaxDD) are in percentage terms. The sample period is from January 31, 2001 to June 30, 2021.

For Commodities, we consider the four broad sectors Precious and Industrial Metals (PM and IM), and Energy and Agriculture (Ags). During the sample period, PM and IM returned 7.51% and 4.57% annualized excess returns at volatilities of 19.30% and 21.63%; Energy and Ags, lost -6.21% and -0.09% p.a. Therefore, while commodity returns are very volatile, the risk is not always rewarded. In contrast, the three commodity style factors, Carry, Quality, and Value, have positive returns, ranging from 2.74% (Carry) to 3.92% (Quality), at considerably reduced volatility. Commodity Quality emerges with the highest Sharpe ratio of 1.45. Only the Momentum factor is negative on average, at -1.06% p.a.

For currencies, we use two baskets: A Developed Market (MSCI EAFE Currency) index and an Emerging Market (MSCI Emerging Markets Currency) index. The former measures the total return of currencies for twenty-one developed countries excluding the U.S. and Canada; the latter measures the total returns for twenty-seven emerging markets. These baskets had average returns of 1.00% and 2.70% p.a. with volatilities of 7.02% and 6.36%, respectively. We also include three FX style factors: Carry, Value, and Momentum. These style factors have enjoyed annual returns of between 1.07% (Value) and 3.34% (Carry).

Inspecting the correlation structure over our entire sample period, we note in Figure 3.1 that global equities are highly positively correlated with Cyclical-Defensives (0.61), the three credit assets (ranging from 0.69 to 0.76), commodities (0.11 for PM to 0.53 for IM), and

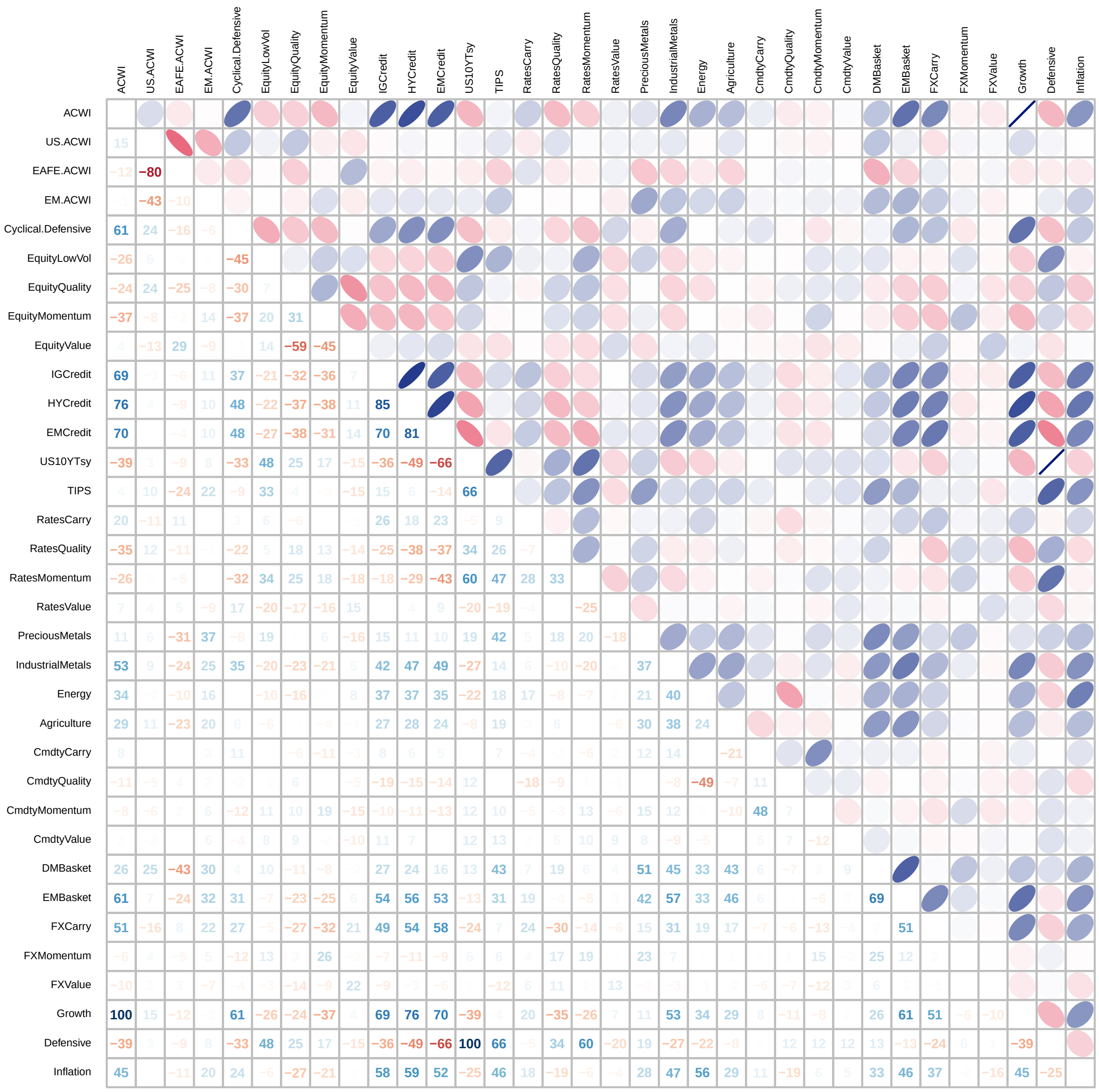


Figure 3.1: Correlation Matrix for the Multi-Asset Multi-Factor Universe. This figure depicts the correlation structure of the multi-asset multi-factor universe, building on monthly data for the full sample period from January 31, 2001 – June 30, 2021. Correlation coefficients below the diagonal are in percentage terms. Colors range from dark red (correlation of -1) to dark blue (correlation of 1).

the EM basket (0.61) and FX Carry (0.51). At the same time, global equities are clearly negatively correlated with Treasuries and equity style factors (-0.26 for LowVol to -0.37 for Momentum). As a result, the three credit assets feature similar correlation characteristics.

Conversely, Treasury correlations range from credits (-0.66 to -0.36) to Equity LowVol

(0.48), Rates Momentum (0.60), and TIPS (0.66). Treasuries are fairly uncorrelated (or even slightly negatively so) with commodities and currencies in general. Rates Carry, Quality, and Momentum are positively correlated, while Rates Value is negatively correlated with its style factor counterparts.

The commodity assets are positively correlated with TIPS (0.14 for IM to 0.42 for PM). They also correlate with the two currency baskets (ranging from 0.33 to 0.57). Interestingly, commodity style factors are almost uncorrelated with most other assets and style factors. Exceptions are the relations between Quality (-0.49 with Energy) and Momentum and Carry (0.48).

Defining macro factors

As stated in Equation (3.1), the above asset class and style factor returns are expressed as a linear function of our macro factors, *Growth*, *Inflation*, and *Defensive*. These allow us to describe, model, and navigate distinct economic regimes and scenarios. Growth and Inflation are of the utmost importance, since they speak to investors' core concerns about expected future cash flows. Naturally, positive growth in an expansion leads to an increase in expected future cash flows, while the opposite holds in a recession. On the other hand, higher inflation leads to a reduction in the present value of expected cash flows.

In addition to Growth and Inflation, we could consider several additional variables, especially if we are concerned about capturing a high proportion of asset return variation (see, for example, Bass, Gladstone, and Ang, 2017). However, we believe it is more practical to focus on a parsimonious set of macro factors to help navigate economic cycles. Therefore, we add a Defensive factor that does well when growth and inflation assets are expected to do poorly.

To further rationalize our choice of macro factors, we conduct a statistical clustering of the multi-asset multi-factor universe. We plot a dendrogram derived from a return correlation-based distance metric (see Lohre, Rother, and Schäfer, 2020).[2] This exercise complements our earlier gauging of correlations, and outlines four clusters (see Figure 3.2). Of these four,

[2]A dendrogram plots objects in a tree, where – moving upward – similar objects are combined into branches. Thus, the higher the height of the fusion, the less similar are the objects.

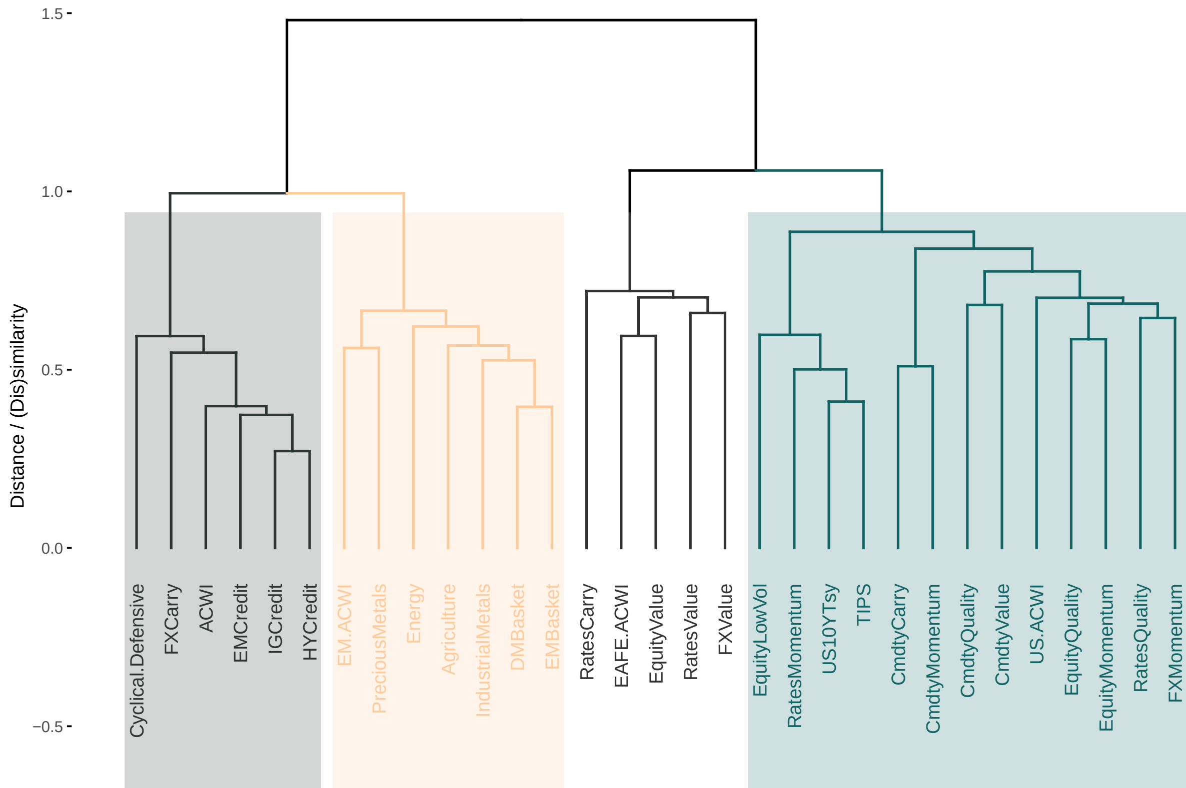


Figure 3.2: Dendrogram for the Multi-Asset Multi-Factor Universe. This figure depicts the dendrogram based on Ward's (1963) method for the multi-asset multi-factor universe, building on monthly data for the full sample period from January 31, 2001 – June 30, 2021.

three are clearly associated with our choice of macro factors. The Growth cluster combines world equity exposure with the three credit spread asset classes and the cyclicals-defensive spread. Given the high correlation of inflation assets with equities in our sample period, the Inflation cluster appears fairly close to the Growth cluster, consisting of all four commodity sectors and the two FX baskets alongside the Equity EM spread.

The Defensive cluster is the largest. We identify the core defensive assets, U.S. Treasury and TIPS, first combining with the truly defensive style factors, Equity Low Volatility and Rates Momentum. The rest of the cluster consists mostly of style factors that are also believed to exhibit defensive properties. Finally, the fourth cluster combines style factors that have a Growth (or rather a non-defensive) character, including Rates Carry and Equity Value.

To inform and anchor the set of asset class and style factor returns, we must choose the best representative for each dimension. The factor Growth is proxied for by the world equity index return, as given by the MSCI ACWI. The factor Defensive is represented by long-term U.S. Treasury bonds. The factor Inflation is measured as TIPS minus U.S. Treasuries, i.e., it gains when inflation rises and declines otherwise. See Table 3.1 for a detailed overview of these factors and their construction.

Our granular choice of macroeconomic factors is connected to earlier work such as Bass, Gladstone, and Ang (2017). They use a principal component analysis to identify the primary return drivers in a multi-asset universe. They document that the first three components (identified as economic growth, real rates, and inflation) account for approximately 85% of the comovements of the selected asset class returns. The additional factors, credit, emerging markets, and commodity, help raise the explanatory power by 10 percentage points. We are thus encouraged to focus on the primary sources of returns – Growth, Inflation, and Defensive – when investigating macro factor allocations. But we also construct an augmented macro risk model with a total of seven macro factors to scrutinize the resulting allocations through a broader macro factor lens.

3.3 From asset to factor diversification

3.3.1 Orthogonal sources of macroeconomic risk

Consider a portfolio of N investable assets with returns $\boldsymbol{R} \in \mathbb{R}^{N \times 1}$. The weighted portfolio return $\boldsymbol{R}^{\boldsymbol{w}}$ is then given as:

$$\boldsymbol{R}^{\boldsymbol{w}} = \boldsymbol{w}^{\intercal} \boldsymbol{R}, \tag{3.3}$$

with portfolio weights $\boldsymbol{w} \in \mathbb{R}^{N \times 1}$.

Let $\boldsymbol{\Sigma} \in \mathbb{R}^{N \times N}$ denote the covariance matrix of asset returns. According to the spectral decomposition theorem, this real symmetric matrix can be factorized into a diagonal matrix of its eigenvalues, $\boldsymbol{\Lambda} \in \mathbb{R}^{N \times N}$, and a matrix $\boldsymbol{E}$ containing the orthogonal eigenvectors of $\boldsymbol{\Sigma}$ as:

$$\boldsymbol{\Sigma} = \boldsymbol{E} \boldsymbol{\Lambda} \boldsymbol{E}^{\intercal}, \tag{3.4}$$

where $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, ..., \lambda_N)$ is the diagonal matrix comprising $\boldsymbol{\Sigma}$'s eigenvalues in descending order, i.e., $\lambda_1 \geq ... \geq \lambda_N$. The columns of matrix $\boldsymbol{E}$ represent the corresponding eigenvectors of $\boldsymbol{\Sigma}$, which define a set of N uncorrelated principal portfolios[3] with variance λ_i for $i = 1, ..., N$. While this decomposition is widely used in the principal component analysis literature, it comes with certain caveats. Despite capturing most asset variation, principal portfolios are often hard to interpret, and may therefore be considered as overly statistical factor portfolios. They can lack economic interpretation and be unstable over time because of estimation errors in the calculation of the covariance matrix (see Bernardi, Leippold, and Lohre, 2018).

Therefore, we use Meucci, Santangelo, and Deguest's (2015) alternative approach to orthogonalizing the factor components. Beginning with a K-factor model, with factor returns $\boldsymbol{F} \in \mathbb{R}^{K\times 1}$ to explain asset returns, their methodology expresses portfolio returns $\boldsymbol{R}^{\boldsymbol{w}}$ in terms of uncorrelated factors $\boldsymbol{F}_{\text{orth}}$:

$$\boldsymbol{R}^{\boldsymbol{w}} = \boldsymbol{w}^\intercal \boldsymbol{R} = \boldsymbol{b}^\intercal \boldsymbol{F} = \boldsymbol{b}_{\text{orth}}^\intercal \boldsymbol{F}_{\text{orth}}, \tag{3.5}$$

where $\boldsymbol{b}$, $\boldsymbol{b}_{\text{orth}} \in \mathbb{R}^{K\times 1}$ denote the factor loadings of the corresponding factors $\boldsymbol{F}$, $\boldsymbol{F}_{\text{orth}} \in \mathbb{R}^{K\times 1}$. Using the factorization (3.4) and representation (3.1) of the investable asset returns as $\boldsymbol{R} = \boldsymbol{B} \cdot \boldsymbol{F} + \boldsymbol{\varepsilon}$, we can explore alternative decompositions of covariance matrix $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{\Sigma}_{\boldsymbol{F}}\boldsymbol{B}^\intercal + \boldsymbol{U} \tag{3.6}$$

where $\boldsymbol{U} \in \mathbb{R}^{N\times N}$ is the assets' idiosyncratic risk that cannot be explained by the given factor structure, and $\boldsymbol{\Sigma}_{\mathbf{F}}$ represents the covariance matrix of the original factors. The key contribution of Meucci, Santangelo, and Deguest (2015) is developing a way to construct orthogonal factor model representations with minimal tracking error. In other words, when creating uncorrelated risk factors, the authors seek to mimic the original factors as closely as possible in order to alleviate concerns about stability and economic interpretation. They

[3] This terminology was introduced by Partovi and Caputo (2004).

choose minimum torsion $\boldsymbol{t}_{\text{orth}}$, which minimizes the tracking error to the original factors as:

$$\mathbf{t}_{\text{orth}} = \underset{\text{Cor}(\mathbf{tF})=\mathbf{Id}_K}{\arg\min} \sqrt{\frac{1}{K}\sum_{k=1}^{K} \text{Var}\left(\frac{(\boldsymbol{t}\mathbf{F})_k - \mathbf{F}_k}{\sigma_k^F}\right)}, \tag{3.7}$$

where $\boldsymbol{t} \in \mathbb{R}^{K\times K}$, $\boldsymbol{Id}_K$ represents the K-dimensional identity matrix, and σ_k^F denotes the volatility of factor $\boldsymbol{F}_k$. Hence, Meucci, Santangelo, and Deguest (2015) construct the orthogonal decomposition of $\boldsymbol{F}$ as a linear transformation $\boldsymbol{F}_{\text{orth}} = \boldsymbol{t}_{\text{orth}}\boldsymbol{F}$, with the corresponding covariance matrix $\mathbf{\Sigma}_{\text{orth}} = \mathbf{t}_{\text{orth}}^{\intercal}\mathbf{\Sigma}_{\mathbf{F}}\mathbf{t}_{\text{orth}}$.

Adapting this linear transformation allows us to express the covariance matrix of investable assets in terms of orthogonalized factors as follows:

$$\mathbf{\Sigma} = \mathbf{B}\mathbf{\Sigma}_{\mathbf{F}}\mathbf{B}^{\intercal} + \boldsymbol{U} = \left(\mathbf{t}_{\text{orth}}^{-1}\mathbf{B}^{\intercal}\right)^{\intercal}\mathbf{\Sigma}_{\text{orth}}\mathbf{t}_{\text{orth}}^{-1}\mathbf{B}^{\intercal} + \boldsymbol{U}. \tag{3.8}$$

To derive factor-mimicking portfolios of these orthogonalized macro factors, we use (3.5), and write asset returns in terms of orthogonalized factors and their beta loadings $\boldsymbol{R} = \boldsymbol{B}\boldsymbol{F} + \boldsymbol{\varepsilon} = \boldsymbol{B}_{\text{orth}}\boldsymbol{F}_{\text{orth}} + \boldsymbol{\varepsilon}$. Following Deguest, Martellini, and Meucci (2013), we define the MFMPs by inverting this relation, and obtain the implied factor-mimicking portfolio returns as:

$$\boldsymbol{R}_{FMP} = \boldsymbol{t}_{\text{orth}}\boldsymbol{B}^{-1}\boldsymbol{R}, \tag{3.9}$$

where $\boldsymbol{B}^{-1}$ is the inverse of the original factor beta loadings matrix estimated using the Moore-Penrose inverse.[4]

3.3.2 Diversified macro factor allocation

Having created MFMPs, we can begin constructing dedicated macro factor allocations. In the absence of a prior view of future economic scenarios, a rational allocation may diversify across the three available macro factor dimensions. Investors usually perceive a portfolio to be well-diversified if it is evenly exposed to different salient sources of risk, i.e., a diversified

[4]In mathematics, this refers to the generalization of the inverse of a matrix as described, for example, by Penrose (1955). This generalization is necessary since $\boldsymbol{B} \in \mathbb{R}^{N\times K}$ is a non-square matrix.

macro factor portfolio would ensure portfolio risk is evenly driven by the three macro risk factors.

Early advocates of this risk parity principle are Qian (2006) and Maillard, Roncalli, and Teiletche (2010). We build on the related framework of Meucci (2009) and Meucci, Santangelo, and Deguest (2015), which proposes a risk parity strategy across uncorrelated risk sources to obtain maximum diversification. With this line of reasoning, it is natural to consider a macro factor risk parity strategy along the orthogonalized MFMPs. By construction, this coincides with simply computing an inverse volatility allocation of the uncorrelated factors with weights:

$$\boldsymbol{w}_{\text{orth}} = \boldsymbol{\Sigma}_{\text{orth}}^{-\frac{1}{2}} \qquad \in \mathbb{R}^{K \times 1}, \tag{3.10}$$

and total portfolio variance emerges as the weighted sum of the uncorrelated factor variances,

$$Var\left(\boldsymbol{R}^{w}\right) = \left(\boldsymbol{w}_{\text{orth}}^{2}\right)^{\intercal} \boldsymbol{\sigma}_{\text{orth}}^{2}, \tag{3.11}$$

where $\boldsymbol{\sigma}_{\text{orth}}^{2}$ is the vector of the minimum torsion factors' variances. Normalizing the factors' contributions by portfolio variance yields a diversification distribution as follows:

$$\boldsymbol{\rho} = \frac{\boldsymbol{w}_{\text{orth}}^{2} \odot \boldsymbol{\sigma}_{\text{orth}}^{2}}{Var\left(\boldsymbol{R}^{w}\right)}, \tag{3.12}$$

where $\odot$ is the Hadamard product. Thus, the diversification distribution simply combines the risk factors' contributions to overall portfolio risk.

Meucci (2009) proposes measuring portfolio diversification as the exponential of the entropy of the diversification distribution, since this corresponds to the *effective number of uncorrelated bets* driving portfolio risk:

$$\mathcal{N}_{\text{Ent}} = \exp\left(-\boldsymbol{\rho}^{\intercal} \ln\left(\boldsymbol{\rho}\right)\right). \tag{3.13}$$

Intuitively, a portfolio driven only by a single risk factor would have $\rho_k = 1$ and $\rho_j = 0$ for $j \neq k$. The resulting effective number of uncorrelated bets would be $\mathcal{N}_{\text{Ent}} = 1$. In contrast, a portfolio with equal factor risk contributions ($\rho_k = \frac{1}{K}$ for all k) gives the maximum number of $\mathcal{N}_{\text{Ent}} = K$ effective bets.

Lastly, we posit expressing the macro factor risk parity portfolio weights not only in terms of MFMPs but also underlying assets. To this end, we simply invert the minimum torsion matrix $\boldsymbol{t}_{\text{orth}}$ and the coefficient matrix $\boldsymbol{B}_{\text{orth}}$ to obtain asset weights: $\boldsymbol{w} = \left(\boldsymbol{B}_{\text{orth}}^{-1}\right)^{\intercal} \boldsymbol{w}_{\text{orth}} = \left(\boldsymbol{t}_{\text{orth}}^{\intercal}\boldsymbol{B}^{-1}\right)^{\intercal} \boldsymbol{w}_{\text{orth}}$ where normalization yields:

$$w^* = \frac{\boldsymbol{w}}{\mathbf{1}^{\intercal}\boldsymbol{w}} = \frac{\left(\boldsymbol{t}_{\text{orth}}^{\intercal}\boldsymbol{B}^{-1}\right)^{\intercal} \boldsymbol{\Sigma}_{\text{orth}}^{-\frac{1}{2}}}{\mathbf{1}^{\intercal}\left(\boldsymbol{t}_{\text{orth}}^{\intercal}\boldsymbol{B}^{-1}\right)^{\intercal} \boldsymbol{\Sigma}_{\text{orth}}^{-\frac{1}{2}}}, \tag{3.14}$$

and $\mathbf{1} \in \mathbb{R}^{K\times 1}$ is a vector of Ones.

3.4 Macro factor investing in practice

3.4.1 Macro factor sensitivities

Next, we are eager to explore the macro factor sensitivities of the chosen asset classes and style factors by regressing their returns on the three macro factors, Growth, Inflation, and Defensive. In other words, we estimate the factor loadings matrix, $\boldsymbol{B}$, in Equation (3.1). Table 3.3 presents the corresponding individual linear multivariate regression results. It focuses on coefficients, t-statistics, and adjusted R^2s to enable gauging the magnitude and source of explanatory power, as well as identifying the macro factor characteristics of each asset class and style factor.

The first panel shows equity investments. By construction, the MSCI ACWI is documented as 100% Growth risk. For the differential returns for the three regions, US, EAFE, and EM, we detect a significant and positive Growth sensitivity for the U.S., which is expected to perform better than the MSCI ACWI in positive Growth environments. Conversely, EAFE has a slight negative exposure to Growth and Defensive, while EM exhibits significant positive loadings on the Inflation factor.

From an equity sector perspective, we are not surprised to find a strong and positive Growth sensitivity for cyclical sectors. They tend to do well in a thriving economic environment, while defensive sectors show merit in more challenging environments. In a similar

Table 3.3: Macro Factor Sensitivities of Asset Classes and Style Factors

	Coefficients			t-stats			Adj. R^2
Macro factors	*Growth*	*Defensive*	*Inflation*	*Growth*	*Defensive*	*Inflation*	
Equities							
ACWI	1.00	0.00	0.00	∞	0.11	-2.31	100%
US-ACWI	0.06	0.06	-0.09	3.11	1.74	-0.96	2.9%
EAFE-ACWI	-0.05	-0.11	-0.15	-2.11	-2.79	-1.30	4.0%
EM-ACWI	-0.05	0.15	0.94	-1.12	1.86	4.00	5.9%
Cyclical-Defensive	0.47	-0.20	-0.25	9.73	-2.31	-1.01	37.1%
Equity LowVol	-0.04	0.31	0.22	-1.80	7.98	1.92	24.8%
Equity Quality	-0.01	0.11	-0.21	-0.74	3.29	-2.28	9.4%
Equity Momentum	-0.16	0.08	-0.06	-4.25	1.21	-0.32	11.2%
Equity Value	-0.01	-0.08	-0.04	-0.34	-2.63	-0.42	1.7%
Fixed Income							
US10YTsy	0.00	1.00	0.00	-6.44	∞	1.21	100.0%
TIPS	0.03	0.67	1.36	3.12	33.62	23.85	86.7%
IG Credit	0.19	-0.07	0.65	10.28	-2.04	6.92	57.0%
HY Credit	0.41	-0.28	1.21	13.24	-5.07	7.53	69.2%
EM Credit	0.37	-0.76	1.02	10.46	-11.92	5.56	70.2%
Rates Carry	0.05	0.03	0.17	2.46	0.93	1.70	4.8%
Rates Quality	-0.05	0.10	-0.02	-3.67	3.82	-0.28	16.3%
Rates Momentum	-0.02	0.37	0.22	-1.29	11.32	2.34	38.1%
Rates Value	0.01	-0.11	-0.16	0.47	-3.17	-1.57	3.9%
Commodities							
Precious Metals	0.12	0.81	2.29	1.29	4.83	4.74	15.3%
Industrial Metals	0.57	-0.17	2.34	6.18	-1.05	4.94	34.5%
Energy	0.13	-0.38	5.46	1.00	-1.64	8.24	30.4%
Agriculture	0.29	0.10	1.49	2.89	0.56	2.93	10.2%
Cmdty Carry	0.03	0.06	0.33	0.76	0.81	1.51	0.7%
Cmdty Quality	0.01	0.06	-0.12	0.72	2.25	-1.57	2.3%
Cmdty Momentum	-0.07	0.14	0.45	-1.32	1.52	1.73	1.4%
Cmdty Value	0.03	0.18	0.22	0.81	2.60	1.09	1.9%
Currencies							
DM Basket	0.11	0.28	0.76	3.43	4.65	4.42	18.6%
EM Basket	0.25	0.13	0.59	9.88	2.94	4.54	43.7%
FX Carry	0.19	-0.02	0.43	6.78	-0.32	3.00	29.0%
FX Momentum	-0.01	0.04	0.05	-0.55	0.90	0.40	-0.6%
FX Value	-0.01	-0.02	-0.23	-0.55	-0.37	-1.89	1.1%

This table reports estimates from linear, separate, multivariate OLS regressions of the assets' and style factors' excess returns on the three macro risk factors. Our sample period is from January 31, 2001 to June 30, 2021.

vein, defensive investment styles, such as Low Volatility or Quality style factors, are expected to work better in downside than upside scenarios. The regression results indeed reveal the clear defensive properties of these two style factors.

Momentum investing in equities exhibits the strongest negative Growth sensitivity. Unsurprisingly, Value is the odd one out, lacking defensiveness as a cyclical investment style. Overall, we expect most equity style factors to help mitigate equity market downturns, mirroring the factor completion efforts of Dichtl et al. (2021).

The second panel of Table 3.3 shows the two asset classes most closely linked to the Defensive factor (US 10Y Treasuries) and the Inflation factor (TIPS). Hence, they exhibit high explanatory power. For the three credit asset classes, we find that the three macro factors explain approximately two-thirds of their variation. Unsurprisingly, Investment Grade, High-Yield, and Emerging Market Credit show large and positive Growth factor sensitivities and negative defensiveness, as they are neutralized with respect to maturity risks. All show a positive Inflation risk sensitivity.

In contrast, the explanatory power for rates factors is much smaller, yielding adjusted R^2s of between 3.9% (Value) and 38.1% (Momentum). For Rates Momentum, we clearly document a strong defensive nature. And we observe a negative Growth loading of Rates Quality, of similar importance as a positive Defensive loading. On the other hand, Rates Value shows negative defensiveness, which would imply a more procyclical investment style.

In explaining commodity sectors, one-third of the variation in Energy and Industrial Metals' returns can be explained by macro factors. Both load positively on Inflation. But Industrial Metals has an additional positive Growth sensitivity, and Energy loads negatively on the Defensive factor (albeit not significantly). These are both intuitive outcomes. For Precious Metals, we continue to find 15.3% in adjusted R^2, driven by positive loadings on the Defensive and Inflation factors. Lastly, Agriculture shows moderate positive Growth and Inflation sensitivities, which translate to an adjusted R^2 of 10.2Ḣence, all four commodity sectors have good inflation hedging capabilities.

Turning to commodity style factors, we observe almost no significant relationship to macro factors, leading to a maximum adjusted R^2 of 2.3% for Commodity Quality (driven somewhat by a positive Defensive sensitivity). We also detect moderate positive defensive factor loadings for Commodity Value.

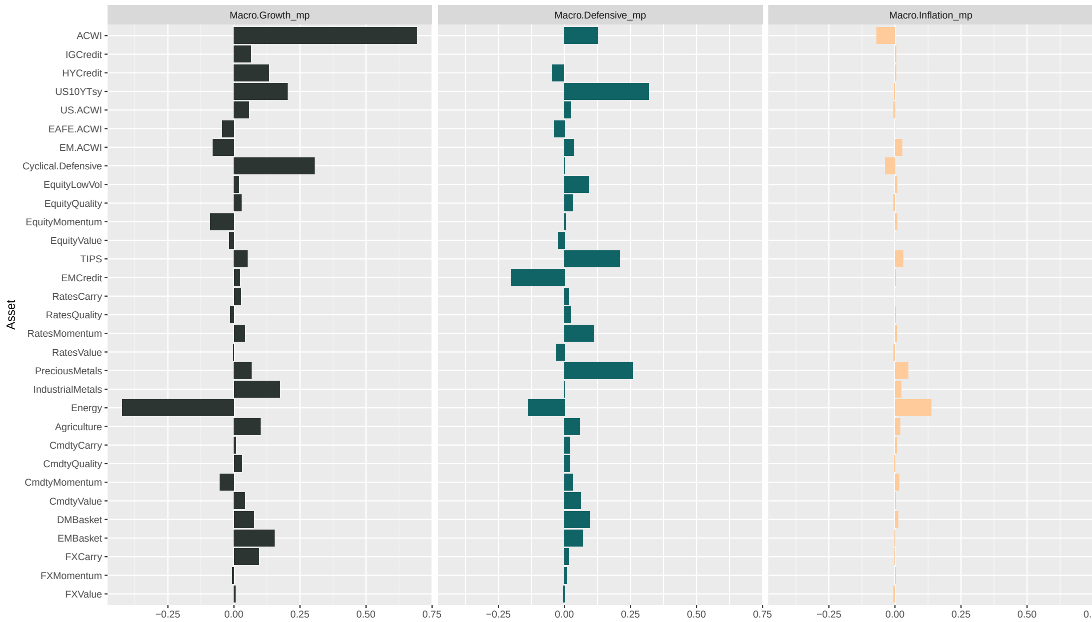
Both the currency baskets, developed and emerging, show similar positive macro factor loadings on all three factors. But the statistical evidence of positive Growth sensitivity is much stronger for EM. Its adjusted R^2 is more than twice as high as that for the DM basket (43.7% vs 18.6%). Similarly to the other asset classes, FX style factors show relatively less explanatory power, but we observe intuitive outcomes. For example, FX Carry exhibits strong positive Growth sensitivity, related to how well it works during calm market periods (but poorer in turbulent environments). In contrast, FX Value and Momentum show negative, although insignificant, Growth sensitivity. Thus, these two style factors may act as diversifying sources to the FX Carry strategy.

3.4.2 Macro factor-mimicking portfolios

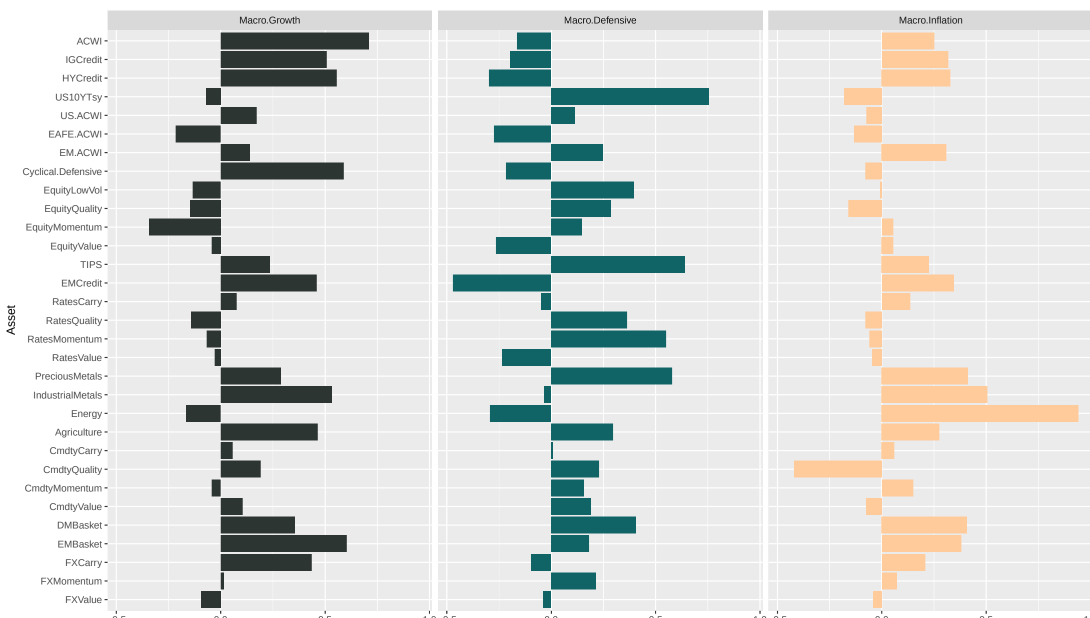
With the macro factor sensitivities for all the asset class and style factors, we can now construct macro-factor mimicking portfolios (MFMPs) according to Equation (3.9). The upper chart in Figure 3.3 depicts the MFMP weights when using full sample data. The lower chart illustrates the MFMP loadings with respect to the underlying asset and style factors, obtained as their correlations with the respective MFMP returns.

The Growth MFMP (first column) comes with a large positive weight in equities and a relatively sizable negative weight in Energy. We can observe positive weights in investment-grade and high-yield credit, and in Cyclicals–Defensives. The Growth MFMP also puts positive weights on the remaining commodity sectors (IM, PM, and Ags). Given their smaller sensitivities, style factors tend to have relatively smaller weights. Unsurprisingly, FX Carry comes with a positive weight, while Equity Momentum and FX Momentum have negative weights. The resulting loadings of assets and style factors on the MFMP are nominally in line with the above positioning. Notably, the large negative Energy weight gives rise to a relatively weak correlation of energy assets with the Growth MFMP.

In contrast, the Defensive MFMP (second column) brings in more explicit style factor exposures. Positive weights in U.S. Treasuries, TIPS, and PM are among the cornerstone holdings of the Defensive basket. In addition, we observe positive weights for defensive style factors such as Equity Low Volatility, Rates Momentum, and Quality, and all Commodity



(a) MFMP Weights



(b) MFMP Loadings

Figure 3.3: Macro Factor-Mimicking Portfolios (MFMPs). This figure highlights the nature of the MFMPs: The upper Panel (a) depicts the full-period asset weights of the three MFMPs. The lower Panel (b) illustrates the loadings of the MFMPs on the underlying asset as measured by the correlation between the mimicking portfolio returns and the respective asset and style factor returns over the full-sample period.

style factors along with positive weights in both currency baskets. There are few negative weights, and EM.Credit and Energy are the most prominent. As a result, the Defensive MFMP is clearly negatively correlated with well-known risk assets or factors, while it correlates positively with defensive ones.

The Inflation MFMP (third column) is characterized by relatively smaller weights. It comes with a negative weight in equities and substantial positive weights across all commodities. It has the highest weights in Energy and PM, concomitant with the Inflation-hedging nature of commodities. As expected, the Inflation MFMP has positive weights in TIPS as well. It seems to be unaffected by style factors in terms of actual weighting, but there are negative loadings for some of the commodity style factors such as Quality and Value. All other loadings are in line with the initial weighting scheme, with a heavy focus on commodities, followed by FX and equities.

Next, we compute macro factor portfolios using an expanding window, with an initial time frame of sixty months. We thus obtain the first out-of-sample observation for February 2006. Figure 3.4 depicts the MFMP decomposition in terms of single asset and factor weights in the left-hand column, and macro factor risk in the right-hand column. The first three rows show individual MFMPs, and the last row shows the macro factor risk parity (MFRP) portfolio. All three MFMPs are capable of mimicking the macro risk in a pure fashion.

Examining the portfolios' macro factor risks shows they load solely on the targeted macro factor. The MFMP for Growth shows clearly leveraged weights during periods of higher volatility (especially during and after the great financial crisis, hereafter GFC), while the weights decrease during periods of steady economic growth, e.g., from 2014 through 2018. The Defensive mimicking portfolio exhibits a smooth weights allocation with little turnover. The MFMP for Inflation highlights the consistent and clearly leveraged investment in commodities, where Energy exhibited a definitive spike during the GFC. However, its weights have been fairly stable since then, resulting in a pure Inflation factor-mimicking portfolio as intended.

Lastly, we evaluate the inverse volatility strategy across the three MFMPs that gives us the maximum diversified MFRP portfolio (see final row in Figure 3.4). By design, the MFRP portfolio risk decomposition views all three macro risk factors as equally contributing over

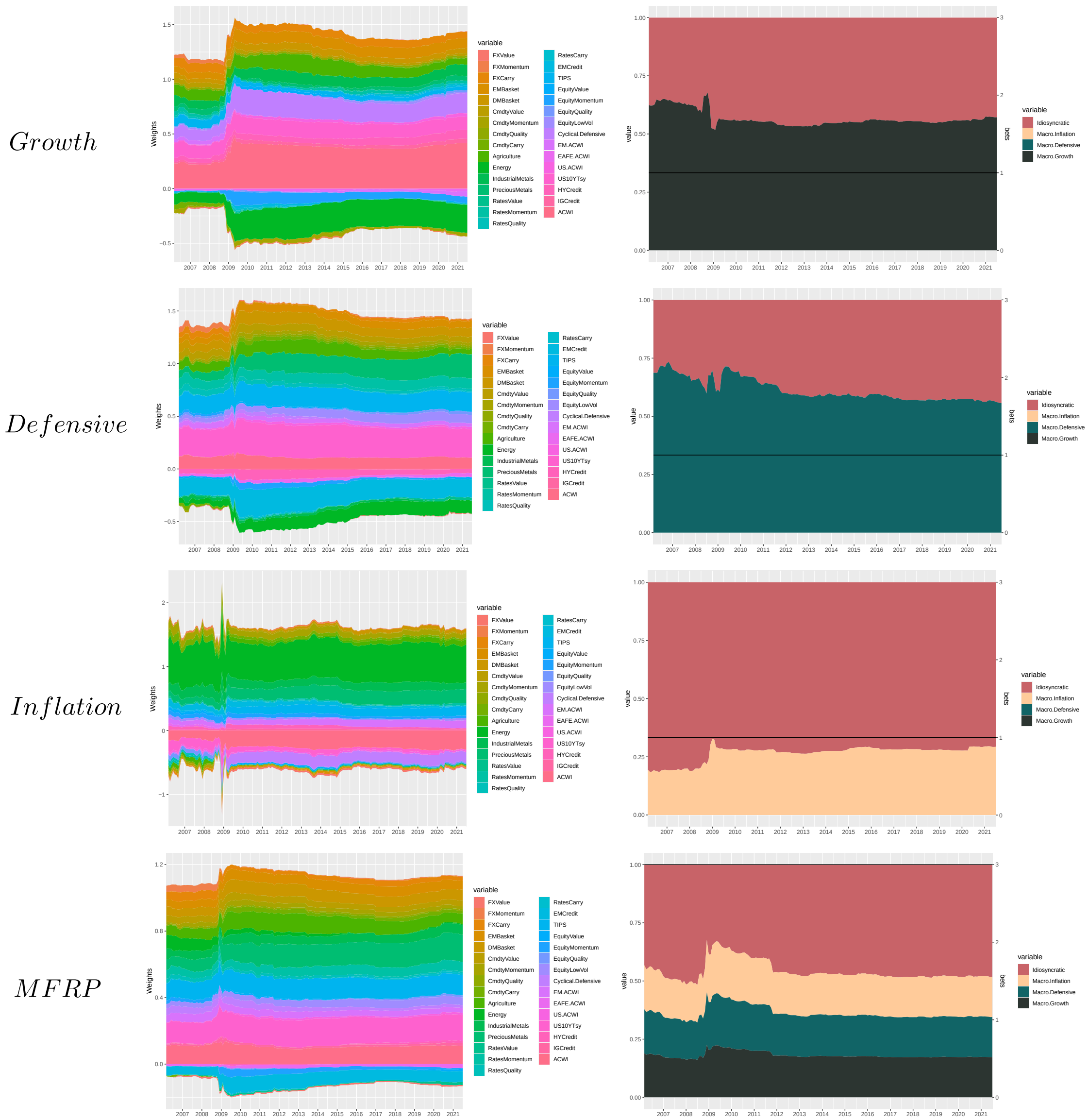


Figure 3.4: MFMP Weights and Risk Decompositions. This figure depicts the decomposition of the macro factor-mimicking portfolios (MFMPs) in terms of single asset and factor weights (left-hand column) and macro factor risk contributions (right-hand column). The results build on expanding window estimations using an initial window of sixty months. The sample period is from January 31, 2006 to June 30, 2021.

time. Importantly, MFRP portfolio weights are fairly stable over time, with little leverage or short-selling activities.

The first four columns of Table 3.4 show the performance statistics of the individual MFMPs, as well as those of the MFRP portfolio. The full period statistics (first panel) illustrate that the Growth MFMP is the most profitable, yielding a historical return of 9.71% p.a. at 8.97% volatility. The Defensive MFMP shows annual returns of 5.98% at

Table 3.4: Performance of Macro Factor Allocations

	MFMP					Macro factor completion		
	Growth	*Def*	*Infl*	*MFRP*	*60/40*	*+Def*	*+Infl*	*+Def+Infl*
Panel 1: Full period								
Net Return p.a.	9.71	5.98	−11.35	3.56	7.34	11.79	7.64	9.95
Volatility p.a.	8.97	7.91	19.70	6.58	10.01	10.54	13.88	12.67
Sharpe Ratio	1.08	0.76	−0.58	0.54	0.73	1.12	0.55	0.79
Max Drawdown	−18.74	−14.51	−90.33	−16.30	−36.54	−22.95	−42.18	−33.74
Calmar Ratio	0.52	0.41	−0.13	0.22	0.20	0.51	0.18	0.29
CVaR	15.66	14.50	43.41	12.46	18.39	18.28	26.13	23.07
Number of Bets	1.00	1.00	1.00	3.00	1.27	2.47	2.30	2.68
Turnover	4.77	4.84	9.12	2.84	0.00	2.62	4.19	5.52
Panel 2: Non-inflationary growth								
Net Return p.a.	19.64	4.07	−28.37	0.42	18.89	17.73	12.03	14.58
Volatility p.a.	4.98	4.06	17.29	4.09	2.88	6.40	6.43	6.38
Sharpe Ratio	3.94	1.00	−1.64	0.10	6.56	2.77	1.87	2.29
Max Drawdown	−1.32	−5.49	−61.73	−7.20	−0.55	−5.03	−5.61	−5.59
Calmar Ratio	14.85	0.74	−0.46	0.06	34.46	3.53	2.14	2.61
CVaR	4.86	7.06	44.18	8.16	0.81	8.18	9.70	8.95
Turnover	3.47	3.77	6.86	1.99	0.00	2.53	5.36	4.30
Panel 3: Inflationary growth								
Net Return p.a.	22.29	1.66	17.36	12.17	31.09	27.13	38.98	35.00
Volatility p.a.	8.98	7.88	15.23	5.93	6.44	9.68	9.63	9.92
Sharpe Ratio	2.48	0.21	1.14	2.05	4.83	2.80	4.05	3.53
Max Drawdown	−3.67	−21.14	−26.99	−6.00	−0.32	−8.53	−4.77	−6.12
Calmar Ratio	6.07	0.08	0.64	2.03	97.88	3.18	8.18	5.72
CVaR	12.46	15.61	26.25	8.78	5.24	12.72	10.00	11.43
Turnover	4.50	4.32	7.24	2.47	0.00	2.38	2.71	5.18
Panel 4: Deflationary and/or crisis								
Net Return p.a.	−9.65	13.32	−33.22	−5.90	−22.75	−8.90	−26.04	−19.05
Volatility p.a.	7.80	9.04	22.21	7.30	9.43	10.82	14.83	13.53
Sharpe Ratio	−1.24	1.46	−1.50	−0.81	−2.41	−0.82	−1.76	−1.41
Max Drawdown	−44.44	−7.55	−89.01	−31.09	−74.79	−40.05	−80.02	−67.61
Calmar Ratio	−0.22	1.76	−0.37	−0.19	−0.30	−0.22	−0.33	−0.28
CVaR	18.79	14.75	56.17	16.62	26.54	24.66	38.61	33.46
Turnover	7.98	8.65	16.54	5.53	2.19	9.19	11.13	11.87

This table shows performance statistics for the macro factor-mimicking portfolios (MFMP) for Growth, Defensive and Inflation, the corresponding MFRP portfolios, the 60/40 portfolio, and the corresponding macro factor overlay portfolios based on either the Inflation MFMP (+Infl), Defensive MFMP (+Def), or both (+Def+Infl). Panels 2-4 highlight the performance of the various portfolios during different regimes. Number of bets denotes the effective number of uncorrelated bets ($\mathcal{N}_{Ent}$). Return, volatility, max drawdown, CVaR, and turnover are in percentage terms. The sample period is January 31, 2006 to June 30, 2021.

7.91% volatility. Its -14.51% maximum drawdown highlights the downside protection of this defensive portfolio, compared to the -18.74% provided by the Growth MFMP at a similar average turnover. Conversely, the inflation MFMP has a negative performance of -11.35%, which highlights the weak performance of inflation assets during our sample period.

Naturally, maintaining a constant inflation hedge as part of the MFRP portfolio induces an associated performance drag in the sample period. We observe a total return of 3.56% at 6.58% annualized volatility. The corresponding Sharpe ratio of 0.54 compares to 1.08 and 0.76 for the Growth and Defensive MFMPs, respectively. Importantly, we note that the MFRP has three effective bets at all times, and a significantly lower turnover (2.84%) than any individual MFMP.

In order to put these results into perspective, we highlight the MFMPs' performance in distinct Growth-Inflation regimes. We divide the sample data into three regimes, using the monthly returns of the simple Growth and Inflation factors as delimiters. The inflationary and non-inflationary growth regimes refer to all months in which the Growth factor rises and the Inflation factor either rises (inflationary) or falls (non-inflationary), respectively. The third regime is characterized by a falling Growth factor.

We find that the various MFMPs perform strongest in their associated regimes. For example, Growth shows extraordinary annual returns in positive Growth environments, but underperforms during periods of negative Growth (19.64% and 22.29% vs. -9.65%). Similarly, Inflation and Defensive succeed in mitigating the macroeconomic risks of inflation and recession, respectively. The Inflation MFMP provides an annual excess return of 17.38% at 15.23% volatility in an inflationary regime. Similarly, Defensive is the only MFMP to show positive performance in a deflationary or crisis regime, with returns of 13.32% and 9.04% volatility p.a. Again, the MFRP portfolio combines the strengths of the three different MFMPs by absorbing most of the downside associated with macroeconomic shocks while giving up a small amount of upside.

3.4.3 Macro Factor Completion Strategies

Scope of diversified macro factor completion strategies

Having explored the diversification benefits of macroeconomic risk factors in an unconstrained set-up, we next examine how we can translate them into fixed traditional multi-asset allocations. In other words, how can we best add to a given allocation to speed macro factor completion?

To illustrate, we choose a static asset allocation, investing 60% in global equities and 40% in bonds (30% in investment-grade corporate bonds, 10% in high yield corporate bonds). First, we assess the ex ante macro factor risk decomposition of this benchmark portfolio. The first row of Figure 3.5 depicts the allocation in terms of asset weights (left column) and macro factor risk contribution (right column). The latter highlights the macro concentration risk of the benchmark portfolio: At any given time during the sample period, more than 85% of the portfolio variance is driven solely by growth risk. This lack of macro factor risk diversification is best summarized by an average number of effective bets of only 1.27 out of the three available bets (Table 3.4, Panel 1, fifth column).

MFMPs for factor completion

Ideally, an investor would choose the MFRP portfolio weights to obtain maximum macro factor diversification. To target this portfolio while adhering to additional constraints, we follow Dichtl et al. (2021). We run a constrained mean-variance optimization that is fed with implied views extracted from the target MFMP (or the entire MFRP). Specifically, we denote the expected returns as $\boldsymbol{\mu} = \gamma \boldsymbol{\Sigma} \boldsymbol{w}^*$, where γ is the risk aversion coefficient, $\boldsymbol{\Sigma}$ is the asset covariance matrix, and $\boldsymbol{w}^*$ is the weights vector of the targeted macro portfolio. We apply a leverage factor of 2 to the MFRP portfolio, and thus $\boldsymbol{\mu}$, because the MFRP needs to match the desired risk level. As a result, the optimization keeps the portfolio weights of the benchmark assets unchanged, but adds the constituents of the respective MFMPs to get as close as possible to the target portfolio. To control turnover and transaction costs, we expand the classic mean-variance optimization problem by a quadratic transaction cost (TC) penalty. We can therefore derive the portfolio weights by solving the following optimization problem:

$$\max_{\boldsymbol{w}} \boldsymbol{w}'\boldsymbol{\mu} - \frac{\gamma}{2}\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w} - \lambda_{TC}\boldsymbol{\Gamma}|\Delta\boldsymbol{w}|^2, \tag{3.15}$$

where $\boldsymbol{\Gamma}$ and λ_{TC} are the asset-specific transaction cost matrix and its scaling parameter, respectively. Using the difference between target weights, $\boldsymbol{w}$, and current holdings, $\boldsymbol{w_0}$, to substitute $\Delta \boldsymbol{w} = \boldsymbol{w} - \boldsymbol{w}_0$ and rearranging terms, the final optimization function yields:

$$\max_{\boldsymbol{w}} \left(\boldsymbol{w}' \left(\boldsymbol{\mu} + 2\lambda_{TC} \boldsymbol{\Gamma} \boldsymbol{w}_0 \right) - \boldsymbol{w}' \left(\frac{\gamma}{2} \boldsymbol{\Sigma} + \lambda_{TC} \boldsymbol{\Gamma} \right) \boldsymbol{w} \right). \tag{3.16}$$

Following Dichtl et al. (2021), we set $\lambda_{TC} = 0.3$, $\gamma = 5$, and assume $\boldsymbol{\Gamma}$ is linear in the diagonal of the variance-covariance matrix.

Figure 3.5 illustrates the diversification effects of adding different MFMPs. The right-hand column depicts the macro factor risk of the benchmark, followed by the outcomes from completing toward Defensive and/or Inflation MFMPs. Note that all three completion strategies work as intended, and show the associated portfolio weight adjustments through time (left-hand column of Figure 3.5).

The considerable Growth exposure of the benchmark portfolio is reduced through the different overlays in all three cases. Adding the Defensive MFMP cuts the Growth exposure in half and increases the number of effective bets by one (from 1.27 to 2.47). We obtain a similar outcome by adding the Inflation MFMP, which reduces Growth exposure by one-third and increases the average number of effective bets to 2.30.

Finally, completing toward the MFRP portfolio provides a stable risk contribution of the three macro factors over time. Growth remains the largest contributor to portfolio risk, but the resulting portfolio is significantly more diversified without overly changing the underlying benchmark.

Rounding out portfolio risk decomposition with macro factor risk contributions is a first-order priority for judging the success of macro factor completion. But we also need to learn about the associated risk-return effects. Considering the Defensive overlay, the downside risk characteristics have changed favorably. Maximum drawdown is now -22.95% (compared to -36.54% for the 60/40 benchmark, see Table 3.4, Panel 1). Moreover, overall return is significantly increased (11.79% vs. 7.34%), while associated volatility is only slightly elevated. As a result, the Defensive overlay brings a considerable increase in risk-adjusted returns.

Conversely, the Inflation hedge would not have added much from a risk-adjusted return perspective. Its maximum drawdown is even more severe (-42.18%). Note that implementing Defensive and Inflation overlays at the same time ultimately results in a middle-ground

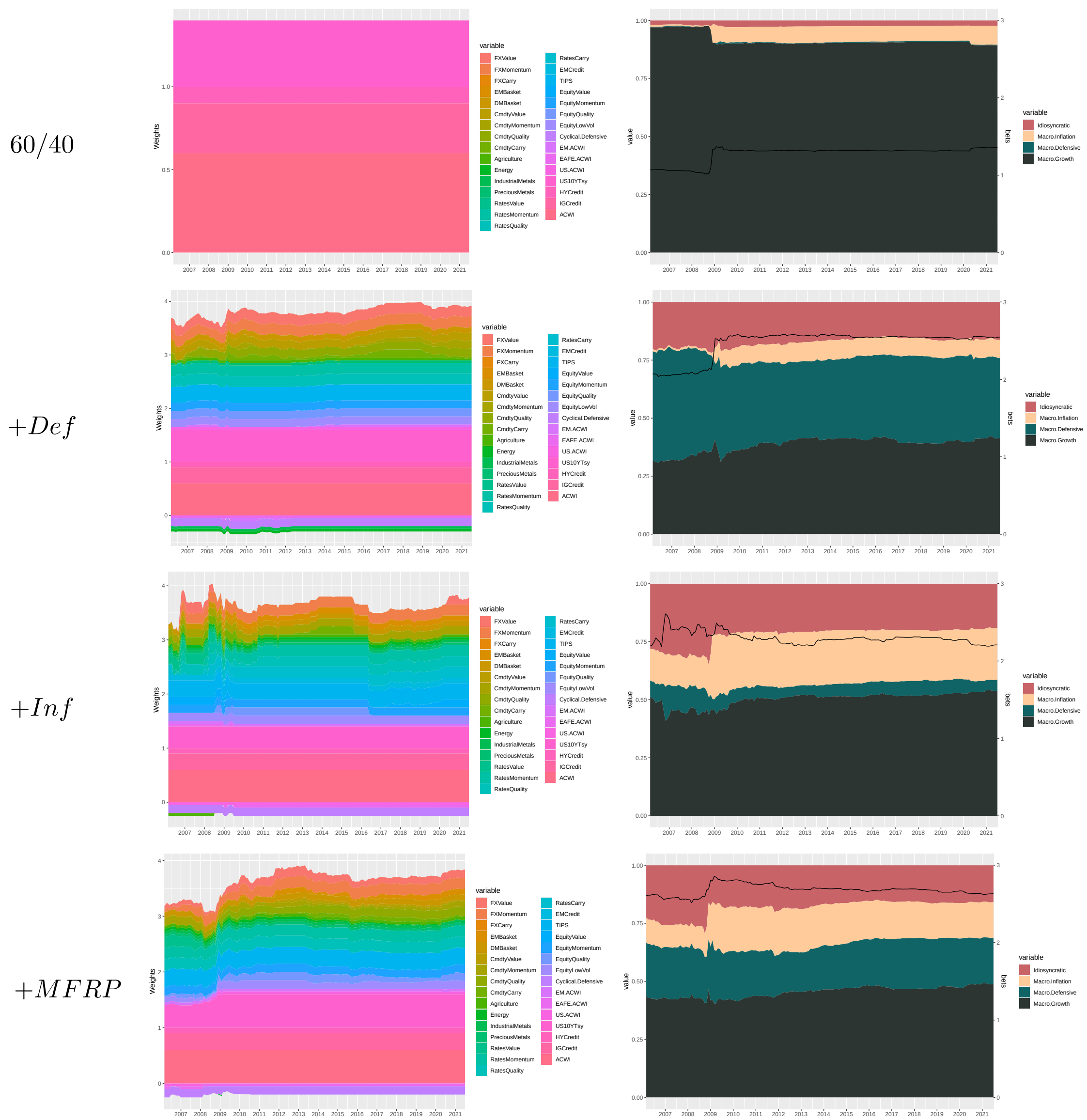


Figure 3.5: Macro Factor Completion Strategies: Weights and Risk Decompositions. This figure depicts the decomposition of the different macro factor completion strategies in terms of single asset and factor weights (left-hand column) and macro factor risk contributions (right-hand column). The results build on expanding window estimations with an initial window of sixty months. The sample period is from January 31, 2006 to June 30, 2021.

position that would have provided some risk mitigation and a pick-up in risk-adjusted return.

Regarding the different regimes, adding the Defensive and/or Inflation MFMP to the benchmark portfolio leads to a significant increase in risk-adjusted performance in the intended environments. Although these overlays decrease risk-adjusted performance in a non-inflationary growth scenario, they serve as effective hedges in less benign regimes (see Panels

2-4, Table 3.4). Adding the Inflation MFMP leads to an increase in performance p.a. of approximately 7 percentage points. In a similar vein, adding the Defensive MFMP to the benchmark portfolio cuts maximum drawdowns during negative Growth regimes by almost half (-40.05% vs. -74.79%), and strengthens portfolio performance from -22.75% to -8.90% for a similar volatility level. Again, adding both Defensive and Inflation MFMPs to the existing portfolio helps mitigate most of the downside associated with the corresponding negative regimes, while preserving growth potential in positive economic growth environments.

Broadening the set of macro factors

In constructing macro factor diversified allocations, we have focused on three factors that relate to economic scenarios investors must navigate: Growth, Inflation, and Defensive. However, we have observed that the associated macro factor-based strategies occasionally come with some portion of risk that cannot be explained by these factors. Hence, to further our understanding of portfolio risk, we augment the macro factor risk model with four factors related to commodity, credit, emerging market, and FX risk. This choice of risk factors is in line with Bass, Gladstone, and Ang (2017), who find that these seven factors amount for 95% of data comovements in a classic multi-asset universe. See Table 3.1 for definitions of the additional macro factors.

Figure 3.6 contrasts the macro factor portfolios' risk decomposition in terms of the three factors used for construction (left-hand column) with that based on the broader lens of seven factors (right-hand column). Generally, the seven-factor model explains a higher fraction of portfolio risk than the three-factor model, as indicated by the lower idiosyncratic risk contribution in the right-hand column of Figure 3.6. It turns out that the Growth and Defensive MFMPs are driven primarily by the corresponding Growth and Defensive factors, respectively. Growth MFMP retains some small contributions from Credit and Commodity, and the Defensive MFMP carries some minor FX exposure in the first half of the sample period. On the other hand, the Inflation MFMP is clearly exposed to Commodity risk as it builds on Precious Metals and Energy assets to weather inflationary periods.

Lastly, the MFRP portfolio — diversifying across Growth, Inflation, and Defensive — shows a balanced profile relative to the broader lens of the seven factors. Given the targeted

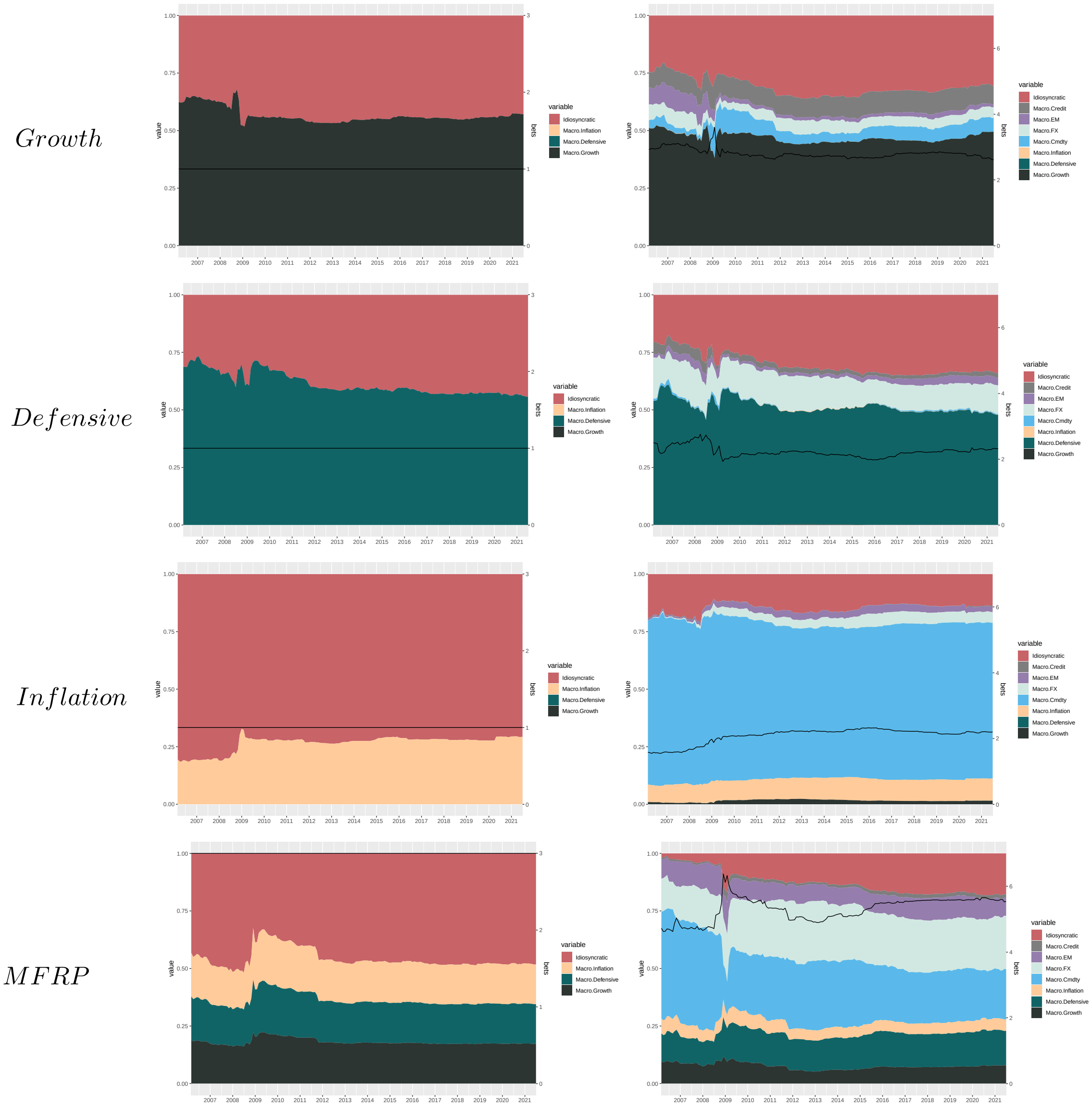


Figure 3.6: MFMP: Seven-Factor Risk Decomposition. This figure depicts the decomposition of the macro factor-mimicking portfolios (MFMPs) in terms of macro factor risk based on a three-factor model (left-hand column) and a seven-factor model (right-hand column). The results build on expanding window estimations with an initial window of sixty months. The sample period is from January 31, 2006 to June 30, 2021.

individual MFMPs, we observe reasonably large risk contributions from defensive, commodity, and FX factors. Nevertheless, this portfolio can also be considered well diversified in terms of the seven factors, giving an average of 5.5 effective bets.

3.5 Discussion

Our proposed framework uses macroeconomic factors to describe economic scenarios and cross-sectional risks. Asset class and style factor returns used in this multi-asset multi-factor analysis have clear sensitivities to the salient macro factors that can be utilized to build out diversified macro factor-mimicking portfolios. Further, we demonstrate how to use these mimicking portfolios in a macro factor completion framework to purposefully manage a given portfolio's macro factor risk exposure.

However, our framework is based on some specific choices. the first important one is the choice of which macro factors to use. We argue to use the three salient macro factors *Growth*, *Inflation*, and *Defensive* as explained in section 3.2. But as outlined in the previous section, these three macro factors come occasionally with some unexplained risks that are not captured by the model. One way to address this concern is to expand the model with additional macro risk factors as described above as well. Naturally, the choice of macro factors has to be based upon either statistically relevant or economically rationalized evidence - or both.

Bass, Gladstone, and Ang (2017) propose a combination of both approaches by decomposing a given investment universe into its statistical risk drivers, i.e., using a principal component analysis to identify salient macro factors. They interpret the statistical factors economically by linking them with traditional asset classes and macroeconomic variables resulting in a set of seven distinct macro factors that account for about 95% of data comovements in the classic multi-asset universe. Our choice of factors might be driven by a slightly different approach of a clustering technique paired with economic rationale, but it leads to very similar results. However, alternative definitions are possible as outlined e.g. in Amato and Lohre (2020).

Another choice we made was the construction of the FMPs via a regression-based approach once the 'original' macro factors were identified. This approach is straightforward and widely accepted in the Finance literature to explain the cross-section of returns. However, when dealing with macro factors this technique might be confronted with a fundamental aspect of macroeconomic data and related factors, i.e., that there is a lot of noise involved due to the lagged nature of published data. Jurczenko and Teiletche (2022) address this concern by proposing a supervised statistical extraction approach that builds upon component selection and target PCA. Their empirical analysis highlights the advantages of such a methodology compared to traditional regression approaches. Hence, we recommend to consider a more advanced method like the one proposed by Jurczenko and Teiletche (2022) to construct even more robust MFMPs.

Pukthuanthong et al. (2019) also address the concern of statistical noise and resulting inefficient FMPs. They compare four different methods proposed by the literature to construct FMPs ranging from cross-sectional regressions to different specifications of using instrumental variables (IV). Their results suggest the use of a two-stage combination of the above where one first constructs the FMPs for each factor via univariate regressions and then applies a multifactor IV Fama-MacBeth regression on those FMPs. This novel approach seems to be more robust to statistical inferences and potential misspecifications of the estimated model.

Overall, our proposed methodology leaves some room for statistical improvements. Nevertheless, this paper aims to highlight the potential use of asset class as well as style factor returns in a combined macro factor framework. Our macro factor framework can thus be used to efficiently steer any given portfolio's macro factor risk exposure in a very pragmatic setup. Future research could analyze how to improve upon this base methodology.

3.6 Conclusion

Traditionally, optimal portfolio allocation has focused on asset classes as building blocks, but this approach can lack sufficient diversification, especially during times of crises. Practitioners have expanded the set of building blocks to include style factors such as Carry, Value, and Momentum. Both asset classes and style factors often have clear sensitivities to macroeconomic factors and associated economic recessions. This prompts us to integrate the allocation of asset classes and style factors into a macro factor investing framework. Specifically, we construct macro factor-mimicking portfolios (MFMPs) to provide diversified vehicles that can better weather recessionary or inflationary periods.

We use MFMPs in a macro factor completion framework to purposefully manage a given mandate's macro factor exposures. More precisely, we demonstrate their optimal positioning given either negative growth and/or rising inflation scenarios. Notably, neither the explicit choice of macro factors nor the specified factor completion anchor portfolio are binding. They can be tailored to the specific investment objective, rendering our approach highly flexible in helping investors diversify macroeconomic risks.

Chapter 4

100 Years of Macro Factor Investing

This project is joint work with my supervisors Sandra Nolte, Mark Shackleton, and Harald Lohre. We thank Guido Baltussen, Mike Chen, Amit Goyal, Clint Howard, and participants at the Robeco Research Seminar in Rotterdam for their helpful comments and suggestions. This work has been supported by an ESRC NWSSDTP CASE Grant.

4.1 Introduction

A guiding principle in the theory and practice of portfolio management is to maximize returns while controlling for associated risks. Growth and inflation risks are amongst a few economic risk drivers that crucially drive investment portfolio performances, and a systematic investor would look to navigate such macro risks through a diversified portfolio. Such an approach rests on a long investment horizon, which eventually will face alternative investment philosophies or styles performing better over shorter horizons. The underlying macro risk premia are rewarded throughout different economic environments, and, therefore, macro factor portfolios come with more balanced risk-return profiles and are expected to be more resilient.

In order to harvest the long-term macro risk premia and exploit their medium-term cyclicality investors can consider dynamically allocating to specific risk premia based on macroeconomic indicators. Such indicators are deemed relevant for explaining time-variation in asset class returns (e.g., Chen, Roll, and Ross, 1986), but less so for explaining the variation of style factors (Ilmanen et al., 2021 or Baltussen, Swinkels, and Van Vliet, 2021). It is thus a challenge to conceive a portfolio allocation that allows investors to diversify across larger macro risks and also to exploit time variation in asset class and style factor returns.

In this paper, we emphasize the use of macro factors for portfolio construction and analyze the robustness of diversified macro factor investing throughout different economic cycles over a 100 year sample period. For augmenting such diversified macro allocation, we investigate a dynamic approach that accounts for the impact of these cycles on the performance of different asset classes and style factors. We thus develop specific macro factor views which result in macro factor weights based on the identified economic cycle. These macro factor views can readily be transferred into tactical allocation decisions at the asset class or style factor level, and we demonstrate how these signals can be accommodated in a Black-Litterman framework.

Our work is related to Amato and Lohre (2020), who analyze macro factor investing based on a broad set of asset classes as well as to Swade et al. (2021) who extend the investment universe to the corresponding style factors. However, the authors focus on rather short samples, leaving them only with few macroeconomic regimes to evaluate. In contrast, our study considerably increases the sample period, allowing us to evaluate multiple economic cycles over a 100-year time period. Therefore, we can take a close look at improving the mapping between macro factor investing strategies and macroeconomic regimes. To this end, we construct a dynamic macro factor allocation strategy in the spirit of Scherer and Apel (2020) that allows adding macro factor views in a Black-Litterman fashion. This methodology helps advance macro factor investing from a strategic allocation perspective towards a macro regime-sensitive investment strategy. Additionally, we demonstrate how to include individual asset class or style factor views such as time-series factor momentum (Gupta and Kelly, 2019).

We contribute to the literature in several ways. First, we extend the macro factor investing evidence by constructing and analyzing macro factor portfolios over a 100-year period. As a result, we can account for different economic regimes and analyze their impact on the robustness of the constructed portfolios. Second, we adapt models used in the business cycle literature to pair the construction process of our macro factor portfolios with identified macroeconomic cycles. The resulting dynamic macro factor allocation enables capturing the cyclicality of macro factors, asset classes, and style factors alike while mitigating salient macroeconomic risks. Third, we incorporate these macro factor views in a classic Black-Litterman framework by transforming them to investable style factor and asset class views. The resulting setup can readily be augmented by further individual style factor views, offering a highly versatile strategic and tactical macro allocation framework.

Our results confirm the medium-term cyclicality of macro-based portfolios and their sensitivity to different economic regimes and periods of distress. We build diversified macro factor mimicking portfolios that retain the genuine macro characteristic but prove to be more robust over the last century than the underlying macroeconomic factor. Exploiting their time variation a tactical allocation overlay yields an information ratio of 0.49 out-of-sample compared to a diversified macro factor risk parity portfolio. Combining the latter with time-series momentum signals increases the information ratio to 1.73. The outperformance is even more pronounced in some economic regimes, with recovery periods benefiting the most.

The remainder of this paper is structured as follows. Section 4.2 reviews factor modeling and investing as well as the challenges associated with investing through different economic cycles. It also explains how to back out macro-factor mimicking portfolio weights based on the orthogonalization technique of Meucci, Santangelo, and Deguest (2015). Section 4.3 navigates a sample of 100 years through the lens of a macro factor investor by constructing robust macro factor-mimicking portfolios based on asset classes and style factors and testing for their diversification properties in different macroeconomic regimes. Following this strategic allocation approach, Section 4.4 emphasizes the use of a dynamic business cycle model to navigate macro factor premia more effectively. The resulting portfolios are complemented by predictive style factor and asset class tilts. Section 4.5 concludes.

4.2 Macro factors and mimicking portfolios

4.2.1 Reviewing macro and style factor research

Ever since the introduction of the seminal capital asset pricing model (CAPM)[1] and Ross's (1976) arbitrage pricing theory (APT), a variety of factor models have been put forward to explain asset returns. Although these models differ in the choice of explanatory factors, they all follow the same intuition. For instance, under APT, the returns $\boldsymbol{R} \in \mathbb{R}^{N\times 1}$ of N risky assets follow a factor intensity structure expressed as:

$$\boldsymbol{R} = \boldsymbol{B} \cdot \boldsymbol{F} + \boldsymbol{\varepsilon}, \tag{4.1}$$

where $\boldsymbol{F} \in \mathbb{R}^{K\times 1}$ represents the returns of K systematic factors with respective factor loadings $\boldsymbol{B} \in \mathbb{R}^{N\times K}$ and asset-specific idiosyncratic risks $\boldsymbol{\varepsilon} \in \mathbb{R}^{N\times 1}$, which are assumed to be uncorrelated across assets and factors and have zero mean.

Following APT, many different factor models have been developed to determine asset prices (see, for example, Fama and French (1993, 2015); or Hou, Xue, and Zhang (2015) in the realm of equity factor models). Most factor models deal with asset-specific factors, i.e., factors that are constructed by sorting on asset class-specific characteristics. Notwithstanding, various studies have shown macroeconomic variables to be relevant in explaining

[1]See Sharpe (1964), Lintner (1965), Mossin (1966) and Treynor (1961).

individual asset prices and even emphasized that asset prices are not only sensitive to economic news but also find the related risks being priced, see, e.g., Chen, Roll, and Ross (1986), Fama and French (1989), Pontiff and Schall (1998), or Ilmanen, Maloney, and Ross (2014). Macroeconomic variables, such as the term spread, industrial production, or inflation shocks, have not only been tested for their effect on asset class returns but also on individual style factors. For example, Chordia and Shivakumar (2002) suggest that profits of momentum strategies can be explained by a set of lagged macro variables, and adjusting for such variables cuts momentum profits significantly.

In a similar vein, Cooper, Mitrache, and Priestley (2022) as well as Kirby (2019) document significant explanatory power of macroeconomic variables for the value and momentum factors. They highlight the time-dependency of different style factor returns with respect to macroeconomic regimes and structural breaks. Yet, although Ilmanen et al. (2021) confirm this significant time variation in risk-adjusted style factor returns over a century of historic data, they conclude them hard to forecast based on macroeconomic variables. This is in line with Baltussen, Swinkels, and Van Vliet (2021) who do not find a significant explanation of style factor premia variation in macro variables in a spanning analysis covering about two centuries.

These intuitively controversial findings might be rationalized through structural breaks that separate different economic regimes. Specific factors might come with differentiated performance patterns across these regimes but seem fairly unaffected by macro changes over the whole sample period. In this context, Ang and Bekaert (2004) propose a regime-switching model characterizing different market regimes in terms of expected returns and conditional volatility. They emphasize the strong performance of regime-shifting investment strategies compared to fixed allocations. This is confirmed by various researchers in the context of factor timing strategies based on different regime classifications (see, e.g., Chousakos and Giamouridis, 2020; Polk, Haghbin, and De Longis, 2020). To this end, Markov-switching models are frequently used, see, e.g., Kritzman, Page, and Turkington (2012) who forecast regimes in market turbulence, inflation, and economic growth. An alternative approach for dealing with different regimes has been put forward by Jurczenko and Teiletche (2018). They propose an alternative to the framework of Black and Litterman (1991, 1992) and ultimately use views generated by macroeconomic regime signals to construct a linear combination of a passive risk-based portfolio and a mean-variance optimized portfolio.

Another strand of the style factor timing literature deals with slower-moving models based on economic regimes. These models typically come with fewer changes between the different regimes, i.e., the necessary criteria to pinpoint a regime switch are more restrictive. For instance, Blin et al. (2021) use a nowcasting procedure to identify business cycles, Van Vliet and Blitz (2011) and Scherer and Apel (2020) use classic financial market variables to classify different cycles. All these papers document significant exposure of some style factor strategies to different economic regimes, which suggest some room for profitable timing of style factors.

4.2.2 Constructing orthogonal macro factor-mimicking portfolios

Macro factors directly follow the factor representation in equation (4.1) and the implications of APT. Similar to the equity factor world, there is no unique macro factor model that outperforms all other models and thus leaves a few choices to the researcher/investor. However, besides the choice of factors itself, there is one additional caveat in dealing with macro factors instead of asset class-specific styles factors: Macro factors lack direct investibility. Whilst style factors are usually represented by long-short portfolios sorted by specific stock characteristics, macro factors are hard to invest in without mapping them to tradable products. Therefore, investors typically resort to mimicking portfolios that have similar properties as the given macro factor but consist of tradable assets.

Dealing with the aforementioned challenges of macro factor investing we resort to a specific set of macro factors. Ideally, the chosen macro factor representation would consist of uncorrelated factors to speed the construction of diversified macro factor portfolios (Swade et al., 2021); yet this is rarely present in macro factors. Given that macroeconomic factors need to be mimicked by investable assets, one might as well resort to orthogonal factor versions instead as suggested by Meucci, Santangelo, and Deguest (2015). Given a K-factor model, with factor returns $\boldsymbol{F} \in \mathbb{R}^{K \times 1}$, their approach expresses portfolio returns $\boldsymbol{R}^{\boldsymbol{w}}$ of a weighted portfolio with portfolio weights $\boldsymbol{w} \in \mathbb{R}^{N \times 1}$ for N investable assets in terms of uncorrelated factors $\boldsymbol{F}_{\text{orth}}$:

$$\boldsymbol{R}^{\boldsymbol{w}} = \boldsymbol{w}^{\intercal}\boldsymbol{R} = \boldsymbol{b}^{\intercal}\boldsymbol{F} = \boldsymbol{b}_{\text{orth}}^{\intercal}\boldsymbol{F}_{\text{orth}}, \tag{4.2}$$

where $\boldsymbol{b}$, $\boldsymbol{b}_{\text{orth}} \in \mathbb{R}^{K\times 1}$ denote the factor loadings of the related factors $\boldsymbol{F}$, $\boldsymbol{F}_{\text{orth}} \in \mathbb{R}^{K\times 1}$. A key component is the minimum torsion matrix $\boldsymbol{t}_{\text{orth}}$ to transform the original factors into uncorrelated ones such that $\boldsymbol{F}_{\text{orth}} = \boldsymbol{t}_{\text{orth}}\boldsymbol{F}$. We follow Meucci, Santangelo, and Deguest (2015) and choose the minimum torsion matrix $\boldsymbol{t}_{\text{orth}}$, which minimizes the tracking error to the original factors as:

$$\mathbf{t}_{\text{orth}} = \underset{\text{Cor}(\mathbf{tF})=\mathbf{Id}_K}{\arg\min} \sqrt{\frac{1}{K}\sum_{k=1}^{K} \text{Var}\left(\frac{(\boldsymbol{t}\mathbf{F})_k - \mathbf{F}_k}{\sigma_k^F}\right)}, \tag{4.3}$$

where $\boldsymbol{t} \in \mathbb{R}^{K\times K}$, $\boldsymbol{Id}_K$ represents the K-dimensional identity matrix, and $\sigma_k^F \in \mathbb{R}$ denotes the volatility of factor $\boldsymbol{F}_k$. To arrive at investable factor portfolios, we can compute macro factor-mimicking portfolio (MFMP) weights as $\boldsymbol{t}_{\text{orth}}\boldsymbol{B}^{-1}$, and their returns are given by:

$$\boldsymbol{R}_{FMP} = \boldsymbol{t}_{\text{orth}}\boldsymbol{B}^{-1}\boldsymbol{R}, \tag{4.4}$$

where $\boldsymbol{B}^{-1} \in \mathbb{R}^{K\times N}$ is the Moore-Penrose inverse of the original factor loadings matrix $\boldsymbol{B}$.[2] Using this framework to construct investable MFMPs we can now turn to an empirical application in the next section.

4.3 100 years of macro factor investing

4.3.1 Data

In implementing a macro-factor-based investment approach, we work with a well-defined global set of assets, style factors, and macro factors. Our sample is based on Baltussen, Swinkels, and Van Vliet (2021) and considers 104 years of monthly data from 31 January 1918 through 31 December 2021. The data set is constructed using financial market prices and macroeconomic series from Bloomberg, Datastream, and the OECD. These series are

[2]We refer to Swade et al. (2021) for more details on the construction and attributes of orthogonal factors.

combined with data from Global Financial Data as well as monthly commodity future data from Chicago Board of Trade annual reports.[3] All returns are in excess of local risk-free rates and expressed in U.S. dollars.

Our final set of investable assets features one global index instrument for each of the three asset classes equities, bonds, and commodities as well as four style factors within each of these asset classes and within currencies. The four style factors are betting against beta (BAB), carry, momentum, and value. These style factors are defined as follows: BAB is defined as low beta assets minus high beta assets with positions neutralized for the ex-ante beta, where beta is measured relative to the global asset class portfolio. Carry is defined as the implied yield on each instrument, i.e., futures implied excess dividend yield for equities, the interest rate differential for currencies, excess bond yield plus rolling curve for bonds, and futures implied convenience yield for commodities. Momentum is the 12-month-1-month excess return. Value is the dividend yield for equities, the real yield for bonds, a five-year reversal in spot prices for commodities, and absolute and relative purchasing power parity for currencies.

Table 1.1 gives the descriptive statistics of the described investment universe. For equities, we consider a global equity index with 9.41% return at 15.07% volatility annualized over the sample period from January 31, 1918, to December 31, 2021. The equity style factors exhibit annualized returns ranging from 2.17% (Value) to 7.11% (Momentum) with volatilities around 10.30%. As for fixed income, the utilized global bond index yields 5.06% return at 3.94% volatility, and the corresponding four style factors show annualized returns ranging from 0.76% (BAB) to 6.50% (Carry) and come at 8.16% (BAB) to 11.35% (Carry) volatility. Commodities are the most volatile asset (18.57%) and have an annualized return of 2.69%. The corresponding style factors have returns between 1.27% (BAB) and 5.70% (Momentum). Lastly, we consider four currency factors with annual returns ranging from 0.32% (Value) to 3.35% (Carry). Overall, 15 out of these 20 style factor strategies have Sharpe ratios significantly greater than zero as indicated by the t-statistics in Table 1.1.

Next, we turn to the choice of macro factors. These macro factors are supposed to

[3] For a detailed overview of how the individual time series are constructed, we refer the reader to Baltussen, Swinkels, and Van Vliet (2021) as well as the corresponding online appendix.

Table 4.1: Descriptive Statistics

	Ret p.a.	Vol p.a.	SR	t-stat	Min	Max	MaxDD
Global Equities	9.41	15.07	0.62	6.37	−34.07	25.43	−70.20
Global Bonds	5.06	3.94	1.28	13.10	−5.87	7.92	−10.71
Commodities	2.69	18.57	0.14	1.48	−21.18	28.55	−93.05
Equity BAB	5.65	10.35	0.55	5.57	−36.01	18.04	−43.57
Equity Carry	5.58	10.28	0.54	5.54	−14.28	12.70	−63.97
Equity Momentum	7.11	10.34	0.69	7.01	−11.50	13.90	−28.47
Equity Value	2.17	10.32	0.21	2.15	−15.31	12.09	−75.20
Rates BAB	0.76	8.16	0.09	0.95	−21.53	19.00	−64.58
Rates Carry	6.50	11.35	0.57	5.84	−55.54	11.09	−62.40
Rates Momentum	1.33	11.22	0.12	1.21	−61.00	12.58	−74.19
Rates Value	3.34	10.97	0.30	3.11	−10.90	55.93	−36.14
Cmdty BAB	1.27	9.95	0.13	1.30	−11.42	17.77	−64.49
Cmdty Carry	3.02	10.47	0.29	2.94	−21.29	19.76	−46.61
Cmdty Momentum	5.70	10.64	0.54	5.46	−11.53	13.60	−54.88
Cmdty Value	3.54	10.50	0.34	3.44	−14.29	12.34	−64.47
FX BAB	0.72	6.62	0.11	1.12	−14.05	29.78	−52.01
FX Carry	3.35	11.39	0.29	3.01	−43.63	14.09	−68.35
FX Momentum	3.05	10.97	0.28	2.84	−44.50	39.20	−52.66
FX Value	0.32	10.78	0.03	0.30	−45.15	19.93	−86.77

The table shows descriptive statistics of excess returns for asset classes and style factors. Min and Max denote the lowest and highest monthly excess return during the sample period. SR is the corresponding Sharpe ratio and t-stat reports the t-statistic for testing the null hypothesis that the SR equals 0. Return, volatility, Min, Max, and Maximum Drawdown (MaxDD) are in percentage terms. Sample period: January 31, 1918 to December 31, 2021.

describe typical investors' concerns and external impacts across regions and asset classes. Specifically, we choose a parsimonious set of three macro factors to describe, model, and navigate distinct economic regimes. The three factors are Growth, Inflation, and Defensive. Growth and Inflation directly address investors' core concerns about expected future cash flows. Whilst the growth factor determines future cash flows, inflation determines its current value. The third factor, Defensive, is expected to do well when the other two factors perform poorly.

Before we are able to construct diversified and robust mimicking portfolios of these three macro factors, we have to select time series proxies for these three dimensions. Intuitively, we choose representative global indices as approximations of the true macro factors, i.e., global equities (Growth), commodities (Inflation) as well as global treasuries (Defensive). Given these time series proxies, we next mimic them via our whole set of asset classes and style factors in order to create highly diversified macro factors. Thus, we have tangible and intuitive representatives for each of the three macro factors that cover the long history of our sample. Of course, the individual choice of macro factors and their representative time series can be altered based on investor preference without any loss of generality of our proposed macro factor investing framework. For example, one could simply stick with the three global indices and use them directly instead of creating orthogonal macro factors and resulting mimicking portfolios. Whilst this approach might be straightforward to implement, it lacks of diversification benefits of additional style factor and asset class combinations.

Our choice is in line with alternatives that look to identify the most dominant components of asset return variance in a given portfolio (see, for example, Bass, Gladstone, and Ang, 2017). Whilst they end up using seven macro factors that account for over 95% of the comovement of asset class returns, our chosen macro factors are related to their first three identified principal components that already account for 85% of the cross-asset movements. However, instead of running principal components methods or similar pure statistical approaches ourselves, we choose macro factors close to investable assets by characterizing macro factors as proxies of global asset baskets. Thus, our choice can be rationalized by the characteristics of the aforementioned methods which might come with high statistical relevance and actuality but typically lack economic interpretation. Conversely, alternative approaches determining different states of the economy from fundamental data, e.g., Chen, Roll, and Ross (1986), are closely linked to fundamental economic data and measures but lack precision due to the discrete, low-frequency data-generating process via surveys and ex-post calculations of output numbers.

4.3.2 Constructing robust MFMPs

Given the set of macro factors and investable assets, we aim to construct robust portfolios clearly mimicking the orthogonalized macro factors throughout various economic cycles. We calculate these MFMPs as stated in Equation (4.4). In addition to the three individual

MFMPs we also construct the macro factor risk parity (MFRP) portfolio, that targets equal risk contribution of the orthogonalized MFMPs. The individual factor contributions $\boldsymbol{\rho} \in \mathbb{R}^{K\times 1}$ to the overall portfolio variance can be derived as:

$$\boldsymbol{\rho} = \frac{\boldsymbol{w}^2_{\text{orth}} \odot \boldsymbol{\sigma}^2_{\text{orth}}}{Var\left(\boldsymbol{R}^w\right)}, \tag{4.5}$$

where $\odot$ is the Hadamard product.[4] To gauge the mimicking portfolios' ability to measure the targeted macro factor exposure we compute the *effective number of uncorrelated bets* based on Meucci's (2009) measure of portfolio diversification:

$$\mathcal{N}_{\text{Ent}} = \exp\left(-\boldsymbol{\rho}^{\intercal}\ln\left(\boldsymbol{\rho}\right)\right). \tag{4.6}$$

The effective number of uncorrelated bets will range from 1 (where the portfolio is entirely driven by a single macro factor, i.e., $\rho_k = 1$ and $\rho_j = 0$ for $j \neq k$) to K (for a portfolio with equal factor risk contributions, i.e., $\rho_k = \frac{1}{K}$ for all k).

Our empirical analysis builds on an expanding window with an initial window size of 48 months allowing for out-of-sample observations starting in January 1923. To mitigate adverse effects from potential estimation biases, we run a constrained mean-variance optimization that targets the unconstrained MFMP in light of a quadratic transaction cost (TC) penalty:

$$\max_{\boldsymbol{w}} \boldsymbol{w}'\boldsymbol{\mu} - \frac{\gamma}{2}\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w} - \lambda_{TC}\boldsymbol{\Gamma}'|\Delta\boldsymbol{w}|^2, \tag{4.7}$$

where $\boldsymbol{\Gamma} \in \mathbb{R}^{N\times 1}$ and $\lambda_{TC} \in \mathbb{R}$ are the asset-specific transaction cost vector and its scaling parameter; $\Delta\boldsymbol{w} = \boldsymbol{w} - \boldsymbol{w}_0$ is the difference between target weights, $\boldsymbol{w}$, and initial portfolio holdings, $\boldsymbol{w_0}$. For targeting the specified MFMP, the optimization is fed with expected returns $\boldsymbol{\mu} = \gamma\boldsymbol{\Sigma}\boldsymbol{w}^*$, where $\gamma \in \mathbb{R}$ is the risk aversion coefficient, $\boldsymbol{\Sigma} \in \mathbb{R}^{N\times N}$ is the asset covariance matrix, and $\boldsymbol{w}^* \in \mathbb{R}^{N\times 1}$ is the weights vector of the targeted macro portfolio. Following Dichtl et al. (2021) and Swade et al. (2021), we set $\lambda_{TC} = 0.3$, $\gamma = 5$, and assume $\boldsymbol{\Gamma}$ is linear in the diagonal of the variance-covariance matrix. Note, that all means and covariance matrices in this paper are calculated using sample means and covariance

[4] See Swade et al. (2021) for details on computing risk contributions of single orthogonal factors.

estimators based on the corresponding analysis period.

Figure 4.1 depicts the resulting asset and style-factor portfolio weights (left column) as well as the risk decomposition in terms of macro factor risk (right column). The first three rows show single MFMPs whilst the last row shows the long-only macro factor risk-parity (MFRP) portfolio. All three single MFMPs display asset return loadings similar to that of the pure (unorthogonalized) factors, yet these asset weights are clearly restricted because of the optimization constraints. In addition, all three mimicking portfolios clearly have some style factor exposures which are fairly stable through time. The defensive and inflation MFMPs faithfully mimic the underlying factor, exhibiting close to pure factor risk exposures to the targeted macro factor. Due to the individual asset constraints the growth MFMP struggles especially to solely load on growth risk alone in the first half of the sample, resulting in some exposure to the other macro factors as well. Lastly, we turn to the diversified risk parity strategy along the three MFMPs, labeled MFRP portfolio (last row). Despite constraints, this portfolio maintains equal risk contributions of all underlying macro factors over time. Also, the corresponding asset weights are fairly stable over time without a high turnover or unduly high leverage.

4.3.3 MFMPs through the cycles

We next explore the MFMPs' sensitivity to different macroeconomic market cycles and we leverage 100 years of data spanning multiple economic and market regimes, covering several bear markets and recessions. To classify market states as regimes, we consider the following approaches. First, we differentiate between 'good' and 'bad' market states as characterized by positive or negative global equity returns. Specifically, bull and bear market periods are classified based on calendar year returns of the global equity return series, and we thus pinpoint 22 bear and 78 bull market years in the out-of-sample period. Second, we determine recessionary versus expansionary periods. A given calendar year is classified as recessionary when it is considered recessionary at least six months by the NBER; otherwise, it is classified as expansionary. The out-of-sample period has 15 recessionary and 85 expansionary years. Lastly, we analyze the MFMPs' performance in distinct growth-inflation regimes. Specifically, we divide the out-of-sample period into four regimes, based on the annual return of

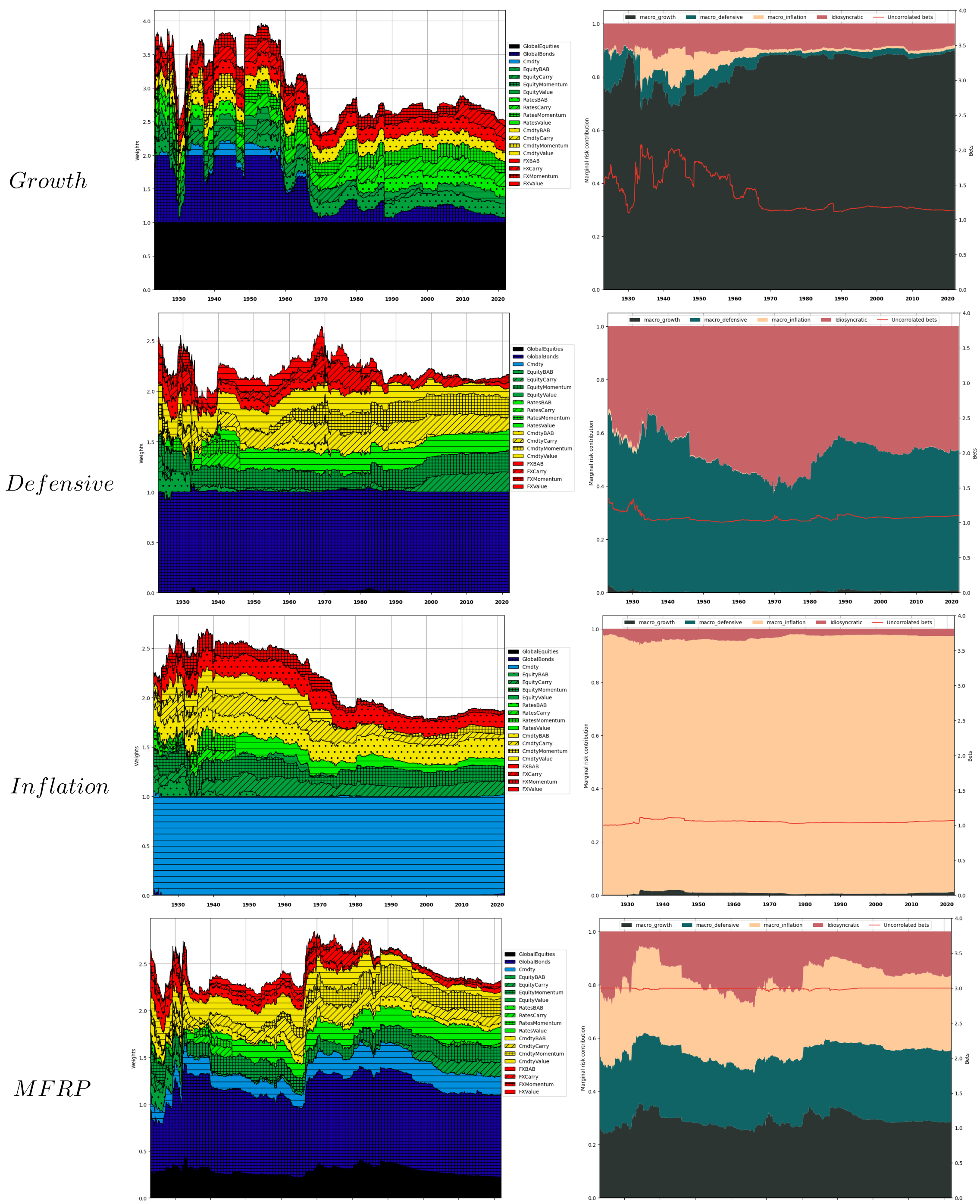


Figure 4.1: MFMP Weights and Risk Decompositions – Long-only. This figure depicts the decomposition of the macro factor-mimicking portfolios (MFMPs) in terms of single asset and factor weights (left-hand column) and macro factor risk contributions (right-hand column) under long-only restrictions. The results build on expanding window estimations using an initial window of sixty months. The sample period is from January 31, 1923 to December 31, 2021.

the simple growth factor and end-of-year real inflation for guidance. The resulting regimes are characterized by positive (negative) annual growth returns in combination with positive (negative) inflation. Our out-of-sample data comprises 6 (2) years of positive (negative) growth and negative inflation as well as 72 (20) years with positive (negative) growth and positive inflation.

Table 4.2 highlights the performance of the three original macro factors (Panel A) as well as the orthogonal MFMPs and the diversified risk parity portfolio MFRP (Panel B). The full sample statistics (first column) illustrate that the growth factor as well as its associated MFMP are the most profitable in terms of absolute returns, yielding annualized historical returns of 10.69% and 19.88%, respectively, followed by the defensive factor with 5.08% and 13.25%, respectively. On a risk-adjusted basis, the latter MFMP is most appealing with a Sharpe ratio of 1.56 compared to 0.98 for the growth MFMP. The inflation MFMP has an annualized return of 10.34% and a Shape ratio of 0.37.

Focusing on the different 'good' and 'bad' states, we clearly see the risks associated with investing in risky assets. Whilst the growth MFMP has extraordinary returns in positive regimes like bull markets (27.93% p.a.) or expansionary regimes (23.93%), its performance lags in 'bad' states (-8.65%) or recessionary periods (-3.06%). The most robust MFMP is the defensive one, which has the highest Sharpe ratios throughout bear versus bull markets (1.61 vs 1.39) as well as recessionary versus expansionary regimes (1.87 vs 1.49).

These observations also hold across the four distinct growth/inflation regimes. Here, the defensive MFMP has the highest Sharpe ratios for the majority of considered macro regimes, i.e. positive growth paired with negative inflation (1.65 for 6 years) or negative growth paired with positive (1.49 for 20 years) or negative inflation (0.38 for 2 years), respectively. However, the growth MFMP performed best with a Sharpe ratio of 1.66 in positive growth and inflation regimes which was the prevailing regime (72 years) in our sample period. The inflation MFMP performed best in the positive growth and inflation regime (0.58), but still trailed behind the other two portfolios.

Overall, we see diversification effects taking place where all MFMPs experience larger risk-adjusted performance measures than the original factors. Yet, we still observe relevant mimicking properties where the MFMPs behave similarly to the macro factors in specific

Table 4.2: Macro Factor Performance in 'Good' and 'Bad' States

			Bull/bear		NBER cycle		Growth/inflation regimes			
Portf	Char	Full	Bull	Bear	Rec	Exp	+/-	+/+	-/-	-/+
Panel A: Original macro factors										
Growth	Ret	10.69	18.01	−15.27	−7.42	13.88	14.16	18.33	−30.12	−13.79
	Std	14.88	13.26	17.67	16.48	14.39	16.63	12.95	17.44	17.67
	Sharpe	0.50	1.12	−1.05	−0.65	0.74	0.66	1.17	−1.91	−0.96
	MaxDD	−58.22	−23.32	−97.76	−83.19	−40.58	−22.32	−23.32	−50.40	−95.74
Defensive	Ret	5.08	5.16	4.83	6.29	4.87	7.74	4.94	−0.62	5.37
	Std	3.91	3.84	4.16	4.36	3.83	3.61	3.85	4.32	4.12
	Sharpe	0.48	0.50	0.39	0.71	0.43	1.25	0.45	−0.89	0.52
	MaxDD	−10.71	−10.18	−12.75	−10.71	−10.18	−1.87	−10.18	−9.98	−5.68
Inflation	Ret	3.39	5.89	−5.49	−11.75	6.06	1.74	6.24	−45.38	−1.50
	Std	18.25	17.68	19.97	20.72	17.68	25.14	16.93	25.32	19.04
	Sharpe	0.01	0.15	−0.44	−0.72	0.16	−0.06	0.18	−1.92	−0.25
	MaxDD	−91.20	−71.26	−82.90	−92.26	−72.60	−60.63	−59.66	−61.30	−66.02
Panel B: Macro factor mimicking portfolios (MFMPs)										
Growth	Ret	19.88	27.93	−8.65	−3.06	23.93	20.24	28.57	−30.88	−6.43
	Std	17.07	15.64	19.26	18.70	16.50	19.90	15.24	19.78	19.13
	Sharpe	0.98	1.58	−0.62	−0.34	1.26	0.86	1.66	−1.72	−0.50
	MaxDD	−58.25	−26.91	−91.02	−74.10	−45.48	−26.91	−26.60	−50.37	−82.49
Defensive	Ret	13.38	13.67	12.34	15.50	13.00	13.62	13.68	5.62	13.01
	Std	6.51	6.50	6.56	6.57	6.49	6.32	6.51	6.28	6.56
	Sharpe	1.56	1.61	1.39	1.87	1.51	1.65	1.61	0.38	1.49
	MaxDD	−11.60	−11.60	−10.59	−7.94	−11.60	−3.94	−11.60	−7.94	−10.59
Inflation	Ret	10.34	13.36	−0.35	−5.05	13.06	8.70	13.75	−42.22	3.84
	Std	19.47	18.97	20.92	21.56	18.99	25.89	18.29	24.42	20.20
	Sharpe	0.37	0.53	−0.17	−0.38	0.52	0.21	0.58	−1.86	0.03
	MaxDD	−85.35	−59.02	−78.15	−86.66	−59.48	−53.07	−51.18	−60.71	−56.47
MFRP	Ret	16.07	18.94	5.91	9.06	17.31	18.26	18.99	−13.69	7.87
	Std	9.22	9.05	9.21	9.93	9.04	10.69	8.91	9.38	9.02
	Sharpe	1.39	1.74	0.29	0.59	1.56	1.41	1.77	−1.80	0.52
	MaxDD	−27.18	−12.90	−38.13	−27.18	−15.46	−9.40	−12.90	−27.18	−17.94
Observations		1200	936	264	180	1020	72	864	24	240

This table shows historical performance characteristics of the original macro factors (Panel A) and long-only MFMPs (Panel B) across various 'good' and 'bad' states based on macroeconomic and market sub-periods based on annual classifications. Return, volatility, and Maximum Drawdown (MaxDD) are in percentage terms. The sample period is from January 31, 1923 to December 31, 2021.

regimes. The single MFMPs are quite robust across various macroeconomic states although the single portfolios perform best in their targeted regime.

Taking full advantage of the diversification effects, we next build the MFRP portfolio amongst the three individual MPMPs. The diversified MFRP portfolio benefits from macro diversification effects absorbing most of the downside associated with poor macroeconomic conditions while only giving up a small portion of the upside. Importantly, the MFRP portfolio has the highest Sharpe ratios for the majority of considered macro regimes, i.e. bull markets (1.74 for 78 years), expansionary periods (1.56 for 85 years) as well as positive growth/inflation regimes (1.77 for 72 years). Nevertheless, it lacks a bit of performance in very specific regimes where especially the defensive MFMP beats the combined portfolio, e.g., bear markets (0.29 vs. 1.39) or recessionary periods (0.59 vs. 1.87). We take a closer look at how to improve the performance of the MFRP during different regimes using a dynamic portfolio allocation in the next section.

4.4 Dynamic macro factor investing

4.4.1 Combining macro and style factor views in a Black-Litterman framework

Section 4.3 suggests a dynamic macro factor allocation looking to benefit from the general macro factor cyclicality might outperform the diversified MFRP portfolio. To this end, we embed tactical allocation signals by complementing strategic portfolio allocations in the Black-Litterman (BL) (1991, 1992) framework. In particular, we generate macro as well as style factor views to refine the expected returns in the mean-variance optimization (4.7). Using the standard master BL formula for refining return and covariance estimates, the return inputs result from:

$$\mu_{BL} = \left[(\tau\Sigma)^{-1} + P^T\Omega^{-1}P\right]^{-1}\left[(\tau\Sigma)^{-1}\Pi + P^T\Omega^{-1}Q\right], \tag{4.8}$$

with Σ referring to the variance-covariance matrix of all assets and style factors as well as $\tau \in \mathbb{R}$ being a scaling constant. $\Omega \in \mathbb{R}^{L\times L}$ represents the view uncertainty by a diagonal covariance matrix of error terms from the expressed views. $P \in \mathbb{R}^{L\times N}$ identifies the assets and style factors subject to views, i.e., it is a projection matrix loading on affected assets and style factors. The equilibrium views are given by matrix $\Pi \in \mathbb{R}^{N \times 1}$ and are backed out from the strategic benchmark (bm) allocation: $\Pi = \gamma \Sigma w_{bm}$. In our case, w_{bm} corresponds to the portfolio weights of the MFRP allocation. The timing signals are collected in the vector $Q \in \mathbb{R}^{L \times 1}$. In our empirical application, we follow Dichtl et al. (2021) and use a prudent specification of Ω and τ such that $\Omega = \text{diag}(\Sigma)$, $\tau = 0.015$. The variance-covariance matrix in equation (4.7) is also adjusted according to the classic Black-Litterman formula[5]:

$$\Sigma_{BL} = \Sigma + \left[(\tau \Sigma)^{-1} + P^{\intercal} \Omega^{-1} P \right]^{-1} \tag{4.9}$$

4.4.2 Macro factor views

In order to time the macro factors, we consider a business cycle model to identify changing macroeconomic environments. Business cycle models are used to categorize the global economic environment by combining various macroeconomic indicators. There is extensive literature about different models and variable choices, and we turn to Van Vliet and Blitz (2011) and Scherer and Apel (2020) for adapting their approaches to our needs. Specifically, we build an aggregated indicator of multiple individual macroeconomic variables to classify the economic state. In order to construct a forward-looking business cycle model we incorporate a combination of market-based indicators as well as output and consumption related ones. Whilst we invest in global assets and factors, we focus on U.S. indicators to enable covering our deep sample.[6]

The used indicators are three fold. First, we incorporate two market-based indicators, the price-earnings ratios of the S&P 500 and the AAA–BAA US spreads. High P/E ratios or small absolute credit spreads indicate a growing economy whilst low P/E ratios or large absolute credit spreads indicate a shrinking one. Second, we use logarithmic changes in production led by 1 month as output related measure. Last, we consider two consumption-related indicators, expected and unexpected inflation, measured as the fitted values and residuals of a full sample regression of inflation data on 12 months of lagged inflation, re-

[5]For a computationally more stable version see Meucci (2010).

[6]The U.S. economy plays a major role for international markets as highlighted for example by Rapach, Strauss, and Zhou (2013).

spectively.[7]

All macroeconomic variables are standardized. The resulting Z-scores are winsorized at three standard deviations using an expanding window, in line with our portfolio construction methodology. The final macroeconomic indicator results as the equal-weighted combination of the Z-scores.

Next, we explain how to use our combined Z-score to classify different regimes. While the state of the economy is captured by the sign of the indicator, the economic trend can be gauged via the sign of the annual change of levels. Combining these two dimensions results in four possible economic regimes which we classify in line with the related literature as Expansion (positive and increasing Z-score), Peak (positive but decreasing Z-score), Recession (negative and decreasing Z-score), as well as Recovery (negative but increasing Z-score). To contain the number of regime switches and related turnover, we either demand two consecutive periods of uniform changes or a significant one-time move that deviates more than one standard deviation from its mean to mark a regime switch.

Figure 4.2 depicts the development of the combined Z-score and resulting business cycle classifications for our sample period from January 1924 to May 2021. Shaded areas indicate NBER recession periods. At first glance, the model captures most of the recession periods and indicates a recovery of a thriving economy otherwise. The model reflects the most extreme shocks during the Great Depression in the 1930s, the early 1980s recession, the Global Financial Crisis in 2008 and 2009, as well as the COVID-19 lockdown in early 2020. These recessionary periods were characterized by extreme stock market turmoil and sell-off. However, the model does not capture all NBER recession periods equally well, especially those that have been milder and/or of rather short duration. For instance, the recessions of 1960/1961, 1969/1970, and the burst of the dot-com bubble in 2001 went unnoticed by our macroeconomic indicator.

Overall, our business cycle model classifies most of the 1,157 sample period months as Expansion (567 months), followed by Peak periods (395 months) as highlighted in Table 4.3. There are only a few periods classified as Recovery (48 months). The average duration of a given regime is 7.85 months, with Expansion periods lasting the longest (9.78 months)

[7] Compare Chen, Roll, and Ross (1986) for a detailed description of the production and inflation measures.

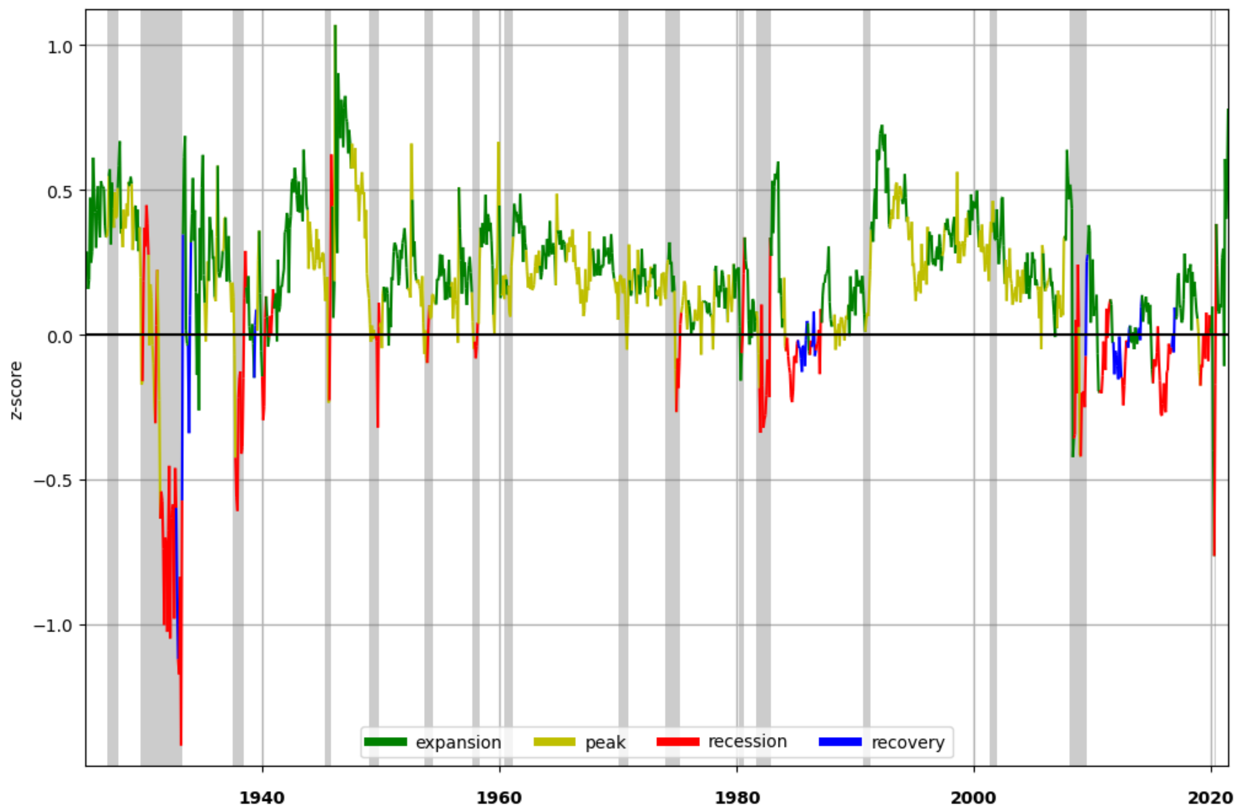


Figure 4.2: Business cycle model. This figure depicts the development of the Z-score of the aggregated business cycle indicator. Shaded areas indicate NBER recession periods. The sample period is from January 31, 1924 to May 31, 2021.

and Recovery periods the shortest (3.69 months). Comparing our model with the NBER recession periods we identify an overlap of 53%. 21% of the Peak periods in our model are categorized as an NBER recession, indicating the slight delay of our market-based indicators compared to the ex-post classification of NBER.

Panel B of Table 4.3 depicts the transition matrix for the four regimes. We clearly document high persistence of regime classifications with probabilities of staying in the same regime as in the previous month ranging from 85% (Recession) to about 90% (Expansion) while the few Recovery periods show lower probabilities of 73%. Conversely, extreme jumps "skipping" one regime in the business cycle have a lower single-digit probability with the highest probability of jumping from a recessionary straight into an expansionary period (6%).

Having thus divided our sample into different business cycle regimes, we next turn to define macro factor view portfolios to inform the BL framework. Our intention in defining

Table 4.3: BCM descriptives

Characteristic	Expansion	Peak	Recession	Recovery	Overall
Panel A: Regime size & NBER overlaps					
Monthly observations	567	395	160	48	1,170
Unique regimes	58	53	25	13	149
Average duration [months]	9.78	7.45	6.40	3.69	7.85
NBER recession overlap [%]	0.06	0.21	0.53	0.19	0.18
Panel B: Transition matrix [%]					
Expansion	89.93	7.95	1.41	0.71	
Peak	10.13	86.58	3.29	0.00	
Recession	6.25	4.38	84.38	5.00	
Recovery	16.67	2.08	8.33	72.92	
Panel C: Annualized return [%]					
Growth	13.92	5.85	6.57	24.85	10.64
Defensive	3.73	5.57	8.50	6.08	5.10
Inflation	7.87	2.96	−10.33	−6.43	3.14
Panel D: Macro view portfolio weights [%]					
Growth	100.00	80.00	80.00	200.00	
Defensive	0.00	20.00	120.00	0.00	
Inflation	0.00	0.00	−100.00	−100.00	

This table shows descriptive statistics of the business cycle model (Panel A), monthly transition probabilities between the different cycles (Panel B), annualized returns of the original macro factors (Panel C), and the resulting macro view portfolio weights (Panel D). The sample period is January 31, 1924 – May 31, 2021.

these view portfolios is not to generate dynamically optimized allocations but rather to select more arbitrary allocations among the three macro factors. Thus, the view portfolios shall express general investors' intuitions on how to best navigate the different regimes from a macro factor perspective.

Therefore, we use the annualized returns in Panel C of the three macro factors and combine them with general investors' intuition to determine the portfolio weights as shown in Panel D. The general line of thought is to participate in the strong upswing potentials of the Growth factor during expansion periods while shifting towards the Defensive factor in peak and recession periods.

In more detail, we clearly see the Growth factor perform exceptionally well in the expansionary period (13.77% p.a.) as well as in the Recovery phase (24.39% p.a.), hence, we want to long the Growth MFMP in these periods. Given the negative performance of the Inflation factor in Recession and Recovery periods, the respective view portfolio is short the corresponding MFMP. Lastly, the Defensive factor clearly outperforms in Recession periods, motivating its large portion in the corresponding macro factor view portfolio. At the same time, we account for potential downturns in the economy during Peak regimes (which still cover 83 months classified as a recession by NBER) without losing too much upside potential by allocating the associated view portfolio a small portion of the Defensive MFMP (20%) and a larger one to the Growth MFMP (80%).

Note that defining portfolio weights based on the proposed way includes a slight look-ahead bias by knowing which factor performs well in which regime. However, we generally do not overfit our results based on any exact or optimized weights but rather follow general intuition on how we allocate across the three macro factors. The resulting extreme portfolio allocations are therefore intentional to incorporate significant views in the BL setup.

Equipped with these macro view portfolios we are able to back out the resulting portfolio weights using equation (4.4) and populate the view vector Q and projection matrix P for the macro factor view case based on the predicted regime in our dynamic macro timing allocation.

4.4.3 Style factor views

In addition to the macro factor views, we also wish to harvest views at the level of individual asset classes or style factors. To this end, we focus on time-series factor momentum, but note that our framework can easily be extended. Moskowitz, Ooi, and Pedersen (2012) demonstrate the possibility to time financial instruments or whole asset classes individually based on their respective performance, given the strong autoregressive structure that these securities typically exhibit. We follow Gupta and Kelly (2019) in constructing a time-series momentum (TSM) strategy to generate asset class and style factor views. The individual TSM return for an individual style factor or asset i with a one-month holding period is given as:

$$f_{i,j,k,t}^{TSM} = s_{i,j,k,t-1} \times f_{i,t},$$
$$\text{where} \quad s_{i,j,k,t-1} = \min\left\{\max\left\{\frac{1}{\sigma_{i,t-1}}\sum_{\tau=k}^{j} f_{i,t-\tau-1}, -2\right\}, 2\right\} \tag{4.10}$$

Equation (4.10) scales the actual factor return $f_{i,t}$ of factor i at time t by the scaling term $s_{i,j,k,t-1}$. That is, the factor gets dynamically scaled by the factor's return standardized by its annualized volatility $\sigma_{i,j,t-1}$ over the formation period j and after an initial exclusion period k. We choose the formation and exclusion periods in line with the related literature and focus on the short-term phenomenon of time series momentum, i.e., $j = 1$ and $k = 0$. The annualized factor volatility $\sigma_{i,t-1}$ is calculated over the previous 36 months. The resulting Z-scores are capped at ± 2.

The combined TSM strategy then combines all individual factor time-series momentum signals into a single long-short portfolio (with formation window j and exclusion window k) as:

$$TSM_{j,k,t} = \underbrace{\frac{\sum_i 1_{\{s_{i,j,k,t-1}>0\}} f_{i,j,k,t}^{TSFM}}{\sum_i 1_{\{s_{i,j,k,t-1}>0\}} s_{i,j,k,t-1}}}_{TSM^{long}} - \underbrace{\frac{\sum_i 1_{\{s_{i,j,k,t-1}\leq 0\}} f_{i,j,k,t}^{TSM}}{\sum_i 1_{\{s_{i,j,k,t-1}\leq 0\}} s_{i,j,k,t-1}}}_{TSM^{short}} \tag{4.11}$$

Hence, the long and short legs are rescaled to form a unit leverage TSM portfolio. We use the corresponding portfolio return as well as the long and short positions to populate the view vector Q and projection matrix P of the BL optimization.

4.4.4 Empirical results

Now, we investigate the efficacy of the tactical overlays based on macro and style factor views. Table 4.4 depicts net performance characteristics of the MFRP anchor portfolio (Panel A), the MFRP portfolio augmented with a tactical overlay based on either a macro signal (Panel B) or a trend signal (Panel C), as well as a combination of both signals (Panel D) over the full sample period but also over the four salient economic regimes. All performance results are accounted for two way turnover[8] and relative performance figures are measured against the MFRP anchor portfolio.

We start our analysis with the MFRP portfolio as a benchmark for following tactical allocation decisions. It comes with an annualized return of 16.04% at 9.17% volatility resulting in a Sharpe ratio of 1.75 for the out-of-sample period from January 1924 to May 2021. The risk-adjusted performance peaks in expansion periods (Sharpe ratio 1.75) and is lowest during Recession periods (1.30). The performance during Recession periods comes with lower returns and simultaneously higher volatility compared to the Expansion periods for example. The differences in performance during different regimes already indicate the potential for improvements when accounting for tactical allocation decisions.

Addressing these timing potentials, we now add our macro factor views to the anchor MFRP allocation (Panel B). One can clearly see an improvement in absolute returns at similar volatility levels for the whole out-of-sample period but also all individual regimes. The resulting Sharpe ratios for the subperiods range from 1.41 (Recession) to 2.66 (Recovery) indicating a clear risk-adjusted improvement as well. Also, the active overlay comes with active returns of 0.72% for the full period at a tracking error of 1.46. The different subperiods yield similar results with the Recovery period being the one where the overlay portfolio deviates the most from its benchmark resulting in an information ratio of 1.23. All other periods experience information ratio between 0.25 (Expansion) and 0.79 (Recession). Despite these positive relative measures, the overall hit ratio, i.e. proportion of outperformance of the overlay compared to its benchmark on a month-to-month basis, is only at 48.50% for the full out-of-sample period. This indicates that the outperformance of the tactical allocation overlay is generated in more extreme months rather than uniformly distributed across the sample.

Next, we analyze the performance of the tactical allocation overlay based on pure trend views. Panel C clearly highlights the strength of the trend signal by boosting the performance of the MFRP portfolio to 22% annualized returns at an almost unchanged volatility level of 9.27% resulting in a Sharpe ratio of 2.39. The active performance increases also to 6.06%

[8]For simplicity, we use trading costs of 10bps for all securities and per trade assuming an efficient execution via futures or similar vehicles.

Table 4.4: Net Performances of Dynamic Factor Portfolios

Characteristic	Overall	Expansion	Peak	Recession	Recovery
Panel A: MFRP anchor portfolios					
Return p.a.	16.04	17.19	14.31	14.83	20.71
Volatility p.a.	9.17	8.76	8.31	11.43	11.79
Sharpe ratio	1.75	1.96	1.72	1.30	1.76
Turnover	2.77	2.34	2.49	4.83	3.32
Panel B: MFRP + macro					
Return p.a.	16.76	17.53	15.05	16.10	23.98
Volatility p.a.	9.18	8.74	8.31	11.39	12.38
Sharpe ratio	1.83	2.01	1.81	1.41	1.94
Active return p.a.	0.72	0.34	0.73	1.27	3.27
Tracking error	1.46	1.35	1.32	1.61	2.66
Information ratio	0.49	0.25	0.55	0.79	1.23
Hit ratio	48.50	46.38	50.89	48.75	53.19
Turnover	22.14	20.12	17.35	32.13	52.79
Panel C: MFRP + trend					
Return p.a.	22.10	22.19	20.77	23.82	26.33
Volatility p.a.	9.27	9.21	8.00	11.44	11.50
Sharpe ratio	2.39	2.41	2.60	2.08	2.29
Active return p.a.	6.06	4.99	6.46	8.99	5.62
Tracking error	3.90	3.23	4.20	5.24	2.96
Information ratio	1.55	1.55	1.54	1.72	1.90
Hit ratio	64.59	63.14	63.80	70.00	70.21
Turnover	71.60	68.44	69.41	87.74	73.30
Panel D: MFRP + macro + trend					
Return p.a.	23.29	23.03	21.94	25.20	31.34
Volatility p.a.	9.30	9.20	8.01	11.25	12.58
Sharpe ratio	2.50	2.50	2.74	2.24	2.49
Active return p.a.	7.25	5.84	7.62	10.37	10.63
Tracking error	4.19	3.55	4.43	5.43	3.68
Information ratio	1.73	1.64	1.72	1.91	2.89
Hit ratio	71.17	68.96	71.39	75.62	80.85
Turnover	79.64	76.92	75.91	93.32	97.09

This table shows net performance statistics for different business cycle-adjusted strategies during different regimes. Relative performance statistics are reported against the MFRP portfolio (Panel A). The anchor portfolio is augmented with business cycle-related macro factor views (+ macro) as well as additional trend views (+ trend). Performances are accounted for two-way turnover. The hit ratio measures the relative months in which the portfolio beats its benchmark. All measures but the Sharpe and information ratios are in percentage terms. The sample period is from January 31, 1924 to May 31, 2021.

although we see a much higher tracking error compared to the MFRP portfolio (3.90%) resulting in an information ratio of 1.55. The performance across the four subsamples comes with a much smaller amplitude implying that the trend overlays work similarly well across all regimes. Notice, that the hot ratio in this setup is far above 50% ranging from 63.14% (Expansion) to 70.21% (Recovery).

Finally combining macro and trend views, we even see an improvement compared to the two distinct overlays. The full out-of-sample Sharpe ratio increases to 2.50 whilst volatility levels stay very similar (9.30%). The resulting information ratio increases to 1.73 with a tracking error of 4.19%. Analyzing the four different regimes we observe more variety: Whilst the Expansion period shows a solid active return of 5.84% to the MFRP benchmark portfolio, especially the Recession and Recovery periods boost the active return with 10.37% and 10.63%, respectively, but also come at slightly higher tracking errors (5.43% and 3.68%). This improved strategy beats its benchmark portfolio in about 70-80% of the time as indicated by the hit ratios for different regimes.

Overall, we find that the additional tactical allocation overlays help to increase the performance of the MFRP benchmark portfolio. Whilst the trend signal boosts the performance across all regimes, we also find the macro factor signal to help improve the performance. This effect is especially strong in periods of a weak economy, i.e., Recession and Recovery periods where we observe the largest benefits of tactical macro factor timing. Thus, the portfolios become more robust in periods of economic downswings while missing less upside potential during economic growth periods.

4.5 Conclusion

We examine the use of macro factors for portfolio construction using a century of data. Therefore, we replicate the three macro factors Growth, Defensive, and Inflation via macro factor mimicking portfolios (MFMPs) and test their persistence throughout different macroeconomic regimes and periods of financial stress. The MFMPs are robust in various economic regimes and hold similar characteristics as the original macro factors. However, through diversification effects they exceed higher risk-adjusted returns and come generally at a lower down-side risk.

We also test the dynamic allocation of macro factors in a Black-Litterman framework. Therefore, we generate macro factor views using a forward-looking business cycle model based on a unified set of macro and market indicators. The resulting z-score is used to classify the state of the economy which then results in a specific macro factor view portfolio. Based upon the resulting views we are able to dynamically over- or underweight the macro factor allocation of our benchmark portfolio. This dynamic allocation results in an information ratio of 0.49 at a tracking error of 1.46 for our full out-of-sample analysis.

Adding additional style factor views in the form of a time-series factor momentum signal helps to further increase the performance of the dynamic portfolio. The information ratio is raised to 0.66 at a tracking error of 1.85 for the full out-of-sample period.

Overall, we show that portfolio construction based on macro factors is viable and works across various economic cycles. Empowering a more conservative risk parity portfolio with active views on macro and style factors even enhances the risk-adjusted performance. These results strengthen the argument that time variation across macro factors are worth to be taken care of although individual style factors might be robust to such macroeconomic changes.

Concluding remarks

In this dissertation, we have delved into the world of systematic investing, shedding light on several key aspects of factor investing. We have analyzed the outperformance of equal-weighted portfolios compared to their value-weighted counterparts, identifying various underlying factors and seasonal effects and emphasizing an alternative approach to harvest the size premium in practice.

Furthermore, we have explored the growing 'factor zoo' by distilling a manageable set of about 15 factors that consistently contribute to the exploitable alpha. Our intuitive and straightforward method is in line with more sophisticated approaches such as double-sorted machine learning techniques, e.g., as in Feng, Giglio, and Xiu (2020). Our findings are not only persistent through time but also robust to different weighting schemes or regions.

Our research also extends its reach beyond equities by venturing into the macroeconomic factor investing domain and introducing robust macro factor-mimicking portfolios designed to navigate the complexities of varying macroeconomic regimes. In order to achieve high diversification, we incorporate not only asset classes but also style factors in the construction of our mimicking portfolios. We analyze the robustness of macro factor investing over 100 years, highlighting the importance of adapting portfolio allocations to match the evolving economic landscape. Our proposed dynamic allocation process generates macro-aware investment portfolios that can outperform classic static benchmark portfolios as seen in significantly positive information ratios.

Overall, this dissertation makes valuable contributions to the field of systematic factor investing. It provides insights into portfolio performance, factors, and macroeconomic dynamics, offering practical guidance and innovative solutions for investors and researchers alike.

Appendix A

Supplementary Research Papers and Articles to Chapter 3

This appendix consists of a research paper that supplements Chapter 3. The research paper is joint work with colleagues from Invesco Quantitative Strategies and has been published in Risk & Reward, Invesco's flagship publication for genuine investment research (Lohre et al., 2020). In this article, we take a macro factor perspective on portfolio allocations and study the link between macro factors and common multi-asset multi-factor investment building blocks. Specifically, we analyze their sensitivities to the two salient macro factors Growth and Inflation, and build robust mimicking portfolios that prove beneficial in diversifying a given portfolio in terms of its macro factor exposure.

A.1 Investing through a macro factor lens

Investing through a macro factor lens

By Dr. Harald Lohre, Scott Hixon, Jay Raol, Ph.D., Alexander Swade, Hua Tao, Ph.D. and Scott Wolle

In brief
A macro factor perspective can help guide portfolio allocation by focusing on salient macroeconomic factors like growth or inflation. We study the link between such macro factors and common multi-asset multi-factor investment building blocks. Specifically, we investigate their macro factor sensitivities and propose a simple, yet effective, route to designing diversified macro factor-mimicking portfolios that prove beneficial in diversifying a given portfolio allocation with respect to its macro factor exposures.

The recent decade has seen a significant rise in factor-based investment propositions, most often focusing on style factor strategies, such as value or momentum. Style factors follow a clear investment rationale and are useful in diversifying a given traditional asset allocation. As many investors are concerned with shocks in macroeconomic variables like growth and inflation, they wish to understand and position multi-asset multi-factor allocations through a relevant macro factor lens.

To set the stage, we briefly recall three general types of factor models as juxtaposed in the seminal paper by Connor (1995). First, *macroeconomic* factor models use macroeconomic variables, e.g. inflation or interest rates, to explain asset returns. Second, *fundamental* factor models use factor portfolio returns related to certain asset characteristics, such as book-to-market or price momentum. Third, *statistical* factor models aim to create factors that naturally hold good explanatory power for the assets under consideration. Yet statistical factor models are often lacking when it comes to shaping the economic intuition of employed factors. Macroeconomic factors, on the other hand, are intuitive but generally provide the lowest explanatory power, leaving a sizable gap of unexplained specific risk.

To strike a balance between these three, we discuss and define macroeconomic factors and investigate the sensitivity of asset classes and style factors with respect to these macro factors. We show why economic regimes matter in constructing effective macro factor-mimicking portfolios and how these can help diversifying macroeconomic risk of a traditional 60/40 asset allocation.

Identifying macro factors
There are generally two distinct approaches for building out macroeconomic factors.[1] The first focuses on pure macroeconomic state variables that can be considered as ultimate drivers of co-movement in asset returns, as in Chen, Roll and Ross (1986). Common macroeconomic state variables are output (to measure growth), inflation, interest rates and risk aversion. However, the explanatory power regarding the returns of many asset classes proves to be modest, complicating the actual implementation of corresponding macro factor-based portfolio allocations.

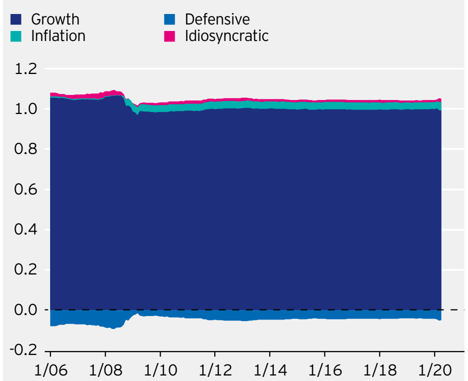
The second approach focuses directly on the factors' ability to explain the cross section of different asset classes' returns. A common statistical methodology to achieve this objective is to run a principal component analysis (PCA) to derive the salient factors explaining most of the asset classes' return variation. In addition, this procedure often creates investable factors. For instance, a PCA typically identifies a portfolio of equities and other risk assets as the most important factor proxying for macroeconomic growth. Similarly, macro factor portfolios representing real rates or inflation risk emerge. Allowing for more granularity in the underlying asset class returns, one may also identify macro risk factors representing commodity, credit, emerging market or currency risk; see Greenberg, Babu and Ang (2016) among others.

> The first factor is growth. The second factor is defensive. The third factor relates to inflation.

To examine the role of macroeconomic factors in portfolio management, we build on the above evidence and focus on three factors in particular: the first factor is growth, as measured by broad equity market exposure. The second factor is defensive, which we proxy by investing in US Treasuries. The third factor relates to inflation and is measured by the spread between inflation-linked bonds and US Treasuries.

Traditional asset allocation through the macro factor lens
To illustrate the relevance of macroeconomic factors, we X-ray a traditional asset allocation in terms of a risk model governed by these three macroeconomic factors. We particularly look into a 60/40 portfolio in global equities and bonds. The 60% equity allocation is represented by the MSCI ACWI index, and the 40% bond allocation splits into 30% in investment grade and 10% high yield bonds. Figure 1 decomposes its

Figure 1
Macro factor risk decomposition of a 60/40 stock-bond portfolio



The chart decomposes the volatility of a 60/40 stock-bond allocation into macroeconomic factor contributions.
Sources: Invesco, Bloomberg, Goldman Sachs.
Sample period: 31 January 2006 to 31 May 2020.

portfolio volatility from 2006 to 2020 into macro factor contributions and shows growth risk to be the biggest (if not sole) contributor to portfolio risk. In the following, we seek to reduce this obvious vulnerability through an allocation process that acknowledges macro factor sensitivities.

Building out diversified macro factor-mimicking portfolios
Asset and style factor data
We wish to investigate the macroeconomic factor sensitivities of a broad set of asset classes and style factors. In each asset class, we aim to be as granular as possible in teasing out the differential element of a given investment. That is, next to broad world equity exposure, we are interested in the returns of certain regions (US, EAFE, EM) relative to the world equity market. Similarly, we look at long-short style factor returns for value, momentum, quality and low volatility investments, isolating the pure factor premia. For fixed income assets, we use US 10Y Treasuries to proxy for the market return and add TIPS, investment grade and high yield corporate bond spreads, as well as emerging market credit spreads. Similar to equities, the factor investing literature supports the notion of fixed income style factors (Kothe, Lohre and Rother, 2021), and we include the four rates factors: quality, value, momentum and carry.

Given the heterogeneity of commodities as an asset class, we abstain from utilizing a broad market index, as these commonly suffer from an extreme energy risk allocation; see Bernardi, Leippold and Lohre (2018) among others. Instead, we investigate the properties of four commodity sectors (precious metals, industrial metals, energy and agriculture) that show little correlation across sectors. We also consider long-short commodity factors along the dimensions carry, value, momentum and quality. Lastly, we include currency investments by allowing two currency baskets, representing the currency allocations implicit in the MSCI EAFE and the MSCI Emerging Markets indices, respectively. We also investigate the three salient currency investment styles carry, value and momentum.

A regime-based route to macro factor-mimicking portfolios

There are different techniques to determine the macroeconomic nature of assets and style factors. For instance, a simple statistical clustering of the multi-asset multi-factor data can help in assembling feasible sets to proxy for a given macroeconomic factor.[2] Another common alternative is to inspect macroeconomic factor sensitivities from multivariate factor regressions. Here, we instead pursue an innovative route that leverages insights from analysis of assets and factors in different economic regimes. As the two decades of multi-asset multi-factor data see a high correlation of growth and inflation assets, we believe such analysis to be vital in identification of genuine growth or inflation assets and factors.

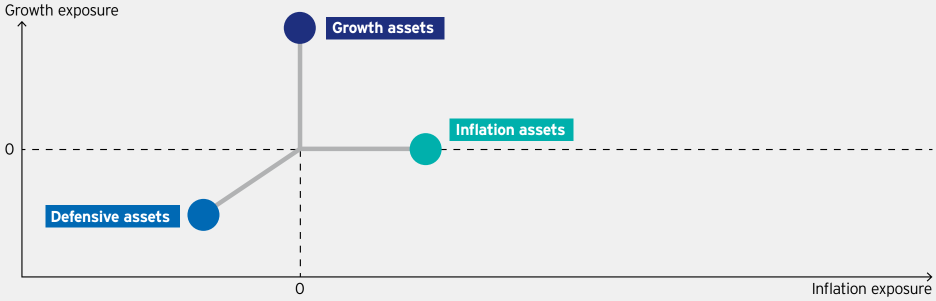
Therefore, to get a sense of how different assets and style factors perform under distinct growth-inflation

Table 1
Determining macro factor-mimicking portfolios

Assets and factors	Rising Growth + Rising Inflation	Rising Growth + Falling Inflation	Falling Growth + Rising Inflation	Falling Growth + Falling Inflation	Growth Exposure	Inflation Exposure	Growth MFMP	Inflation MFMP	Defensive MFMP
Equities									
ACWI	4.77	4.44	-3.91	-3.76	8.44	0.09	2.5%		
US-ACWI	0.77	0.38	-0.60	0.33	0.71	-0.27			16.8%
EAFE-ACWI	-0.74	-0.05	-0.18	-0.35	-0.13	-0.26			18.2%
EM-ACWI	0.60	-0.04	1.23	-0.08	-0.29	0.98		9.3%	
Cyclicals-Defensives	1.58	0.91	-2.03	-1.93	3.22	0.29	11.4%		
Quality	0.59	0.60	0.68	1.92	-0.70	-0.63			29.7%
Momentum	-0.02	0.50	1.30	1.54	-1.18	-0.38			14.9%
Value	-0.51	-0.27	-0.44	-0.86	0.26	0.09			20.4%
Low Volatility	-0.42	0.55	1.50	1.79	-1.57	-0.63			29.8%
Fixed Income									
US 10Y Tsy	-0.75	0.66	1.44	2.17	-1.85	-1.07			30.5%
TIPS	1.10	-0.05	3.42	-0.02	-1.18	2.29		20.9%	
IG Credit	2.25	1.54	-0.38	-2.24	3.21	1.28	16.4%		
HY Credit	2.97	2.18	-0.95	-2.50	4.30	1.16	11.8%		
EM Credit	2.22	1.24	-1.20	-2.64	3.65	1.21	8.5%		
Rates Value	0.44	1.11	-0.12	-0.80	1.24	0.00	12.5%		
Rates Momentum	0.13	0.33	2.18	1.53	-1.63	0.23			17.6%
Rates Quality	-0.48	-0.46	1.81	0.98	-1.86	0.41			26.2%
Rates Carry	0.80	0.93	1.02	0.10	0.31	0.40		13.5%	
Commodities									
Precious Metals	0.90	-0.42	1.04	-0.43	-0.06	1.40		5.5%	
Industrial Metals	1.53	0.45	0.11	-3.28	2.57	2.24		2.8%	
Energy	1.19	-1.08	0.68	-2.97	1.20	2.96		4.7%	
Agriculture	0.69	0.28	-0.44	-1.54	1.47	0.75		3.1%	
Carry	0.81	-0.29	1.09	0.10	-0.33	1.05		9.9%	
Quality	1.25	1.44	1.59	2.20	-0.55	-0.40			32.9%
Momentum	0.14	-0.02	0.97	-0.37	-0.24	0.75		6.5%	
Value	0.73	0.56	0.75	0.62	-0.04	0.15			8.5%
Currencies									
Developed Markets	0.99	-0.58	0.52	-1.35	0.62	1.72		16.8%	
Emerging Markets	2.13	1.05	0.01	-2.71	2.94	1.90	9.8%		
Carry	1.90	1.58	-0.04	-1.14	2.33	0.71	10.7%		
Value	-0.07	0.87	1.08	1.00	-0.64	-0.43			20.8%
Momentum	0.69	0.28	1.35	0.28	-0.33	0.74		15.1%	

The first section of the table gives risk-adjusted performance (Sharpe ratios) of the asset classes and style factors in four different growth-inflation regimes. Columns 5 and 6 synthesize this information into average growth and inflation exposures. The last section of the table shows macro factor-mimicking portfolio weights for the three macro factors: growth, inflation and defensive. For proxies used, please refer to the data appendix at the end of the article.
Sources: Invesco, Bloomberg, Goldman Sachs. Sample period from 31 January 2001 to 31 May 2020. **Past performance is not a guide to future returns.**

Figure 2
Classifying assets and factors based on growth and inflation exposures



Source: Invesco. For illustrative purposes only.

regimes, we divide the sample data in four regimes using the monthly returns of the simple growth and inflation factors as a delimiter. The rising growth/ rising inflation regime comprises all months in which both growth and inflation assets rise. All remaining months fall into the other three regimes (table 1), formed by considering the alternative combinations of growth and inflation returns.[3]

Imagine we expect rising inflation and falling growth. Column 3 shows which assets and factors provided inflation protection in such a regime. Clearly, TIPS stand out with the highest Sharpe ratio. While we would expect good inflation hedge properties for all commodity sectors, this specific regime favors metals and energy. As for the three credit asset classes, we note that their positive correlation to inflation is mostly driven by their proximity to growth assets. However, this relation breaks down in negative growth environments, when credit markets failed to provide inflation protection. Hence, these asset classes clearly can be considered in the growth bucket alongside equity exposure.

Regarding style factors, we observe consistent inflation hedging returns for commodity carry and momentum, as well as FX momentum. To systematically pin down genuine growth, defensive or inflation assets, we use a straightforward procedure to determine the average growth and inflation exposure based on evaluating the differential performance of assets and factors across the various growth-inflation regimes. To illustrate, for an asset (or style factor) to be considered a growth asset, we would expect it to have higher risk-adjusted performance in positive versus negative growth regimes. Specifically, we would wish to observe such outperformance in inflationary and deflationary periods. Hence we define an asset's average growth exposure as the growth spread in risk-adjusted performance averaged across the two inflation regimes. Conversely, inflation assets are expected to do well in inflationary periods, independent of the prevailing growth regime. We thus define an asset's average inflation exposure as the inflation spread in risk-adjusted performance as a simple average across the two possible growth regimes. Table 1 gives these average growth and inflation exposures for all assets and style factors.

With this information, we can plot all assets and style factors on the growth and inflation dimensions. As shown in figure 2, defensive assets would ideally have negative loadings on both growth and inflation; inflationary assets would have zero exposure to growth and large positive loadings on inflation; and growth assets would have zero exposure to inflation and large positive loadings on growth. In practice, many assets will not fit cleanly into one of these three areas, but we at least have clear priors about what constitutes an ideal asset in each macro factor.

We operationalize this idea using simple parameters to define an area for each macro factor. More sophisticated approaches are certainly available, but in this case, we want to illustrate the usefulness of macro factors even with fairly simple definitions. For example, consistent with figure 2, an asset is labeled 'growth' if it is closest to the growth coordinates of the asset or factor with the highest growth exposure. Figure 3 illustrates this procedure using growth and inflation exposures computed over the full sample period. High yield credit has the highest growth exposure and the latter forms the center of the blue area, where growth assets and factors are located.[4] Similarly, energy assets have the highest inflation exposure, forming the center of the purple area, which contains inflation assets and factors. Lastly, the center of the turquoise defensive assets and factors area is determined by the asset with the smallest growth exposure, i.e. 10-year US Treasuries.

Such an approach gives rise to intuitively appealing classifications. For instance, the basket of inflation assets and factors features TIPS, all commodity assets but also a few style factors, such as commodity carry and momentum or FX momentum and rates carry. As for growth, the equity and credit assets are joined by cyclical versus defensive sectors, EM currencies and two style factors, rates value and FX carry, which resonates with the latter suffering in similar periods like equities.

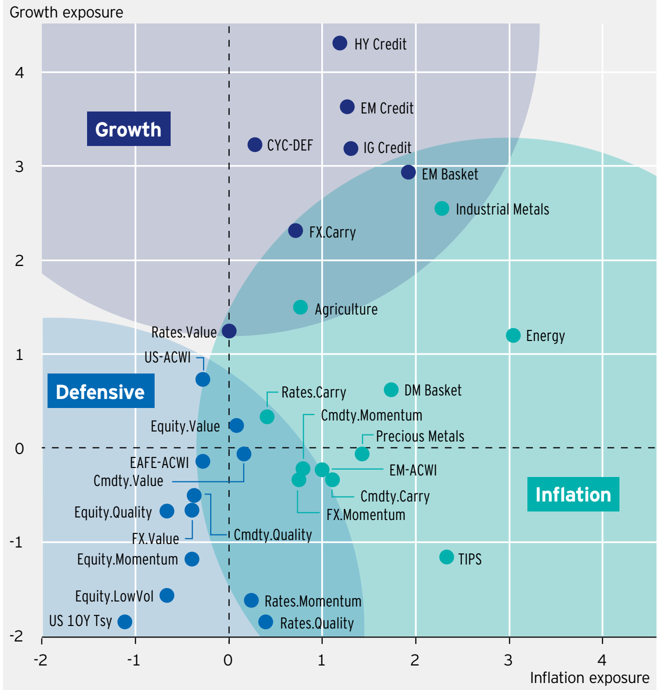
Interestingly, the defensive basket features a larger number of style factors, including almost all quality style factors as well as equity momentum and low volatility. Note that our procedure assigns every asset and factor to one of the three macroeconomic factors.

Based on this macroeconomic classification of asset and style factors, we build three macro factor-mimicking portfolios (MFMPs) that can help guide macro factor-based portfolio allocations. Obviously, the sensitivity of an asset to a certain macro factor decreases with the distance from the respective macro factor's center.

Applying this classification over time through an expanding window, we observe that, while some assets and factors can be clearly associated to one of the three macro factors, others might be reasonably close to more than just one factor. It seems natural to apply less weight to such distant assets and factors when constructing macro factor-mimicking portfolios. Also, we wish to diversify identified macro baskets in terms of risk and therefore apply a straightforward and robust weighting scheme. Specifically, we perform an inverse volatility allocation where the assets' and factors' volatilities are scaled according to their relevance for the macro factor concept. That is, a more distant asset or factor will experience a more severe volatility penalty than a very close one. As a result, the macro factor-mimicking portfolios focus on truly representative assets and factors rather than being unduly dominated by weaker contenders.

The specific constituents and weights for all three MFMPs are shown in the last three columns of table 1; the weights are scaled such that all MFMPs target a volatility of 5%. These portfolios each represent pure exposure to either growth or inflation or defensiveness and thus form meaningful instruments to navigate portfolios through a macro factor lens. Figure 4 shows the macro factor risk decomposition of the defensive and inflation portfolios, suggesting that both MFMPs live up to their respective objective.

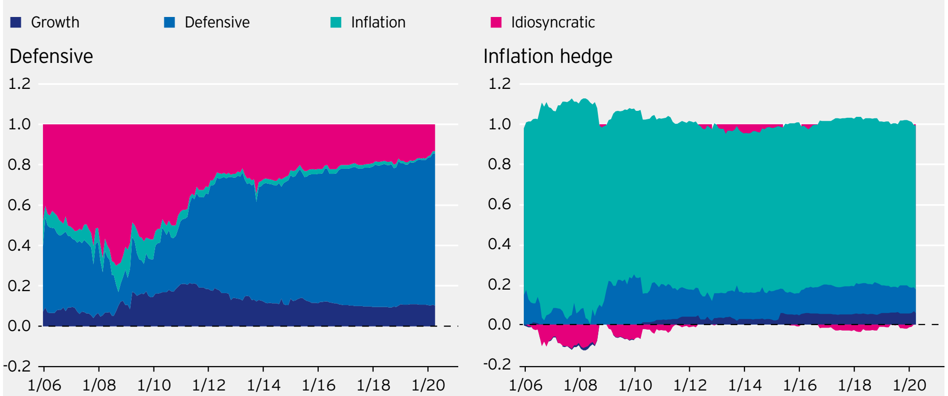
Figure 3
Determining macro factor-mimicking portfolios



The chart depicts asset and style factors in a scatter plot according to their average inflation exposure (x-axis) and average growth exposure (y-axis).
Source: Invesco. Sample period from 31 January 2001 to 31 May 2020.

Macro factor-based portfolio overlays
We now make use of the macro factor-mimicking portfolios. Circling back to the concentrated growth risk allocation of the 60/40 stock-bond allocation, we explore ways of altering the risk profile. First, we add a defensive overlay to reduce growth risk in a 60/40 portfolio. Second, we add an inflation hedge to help protect against an increase in inflation. Third, we consider the effect of combining defensive and inflation hedge overlays with the 60/40 allocation.

Figure 4
Macro factor portfolios – defensive and inflation hedge



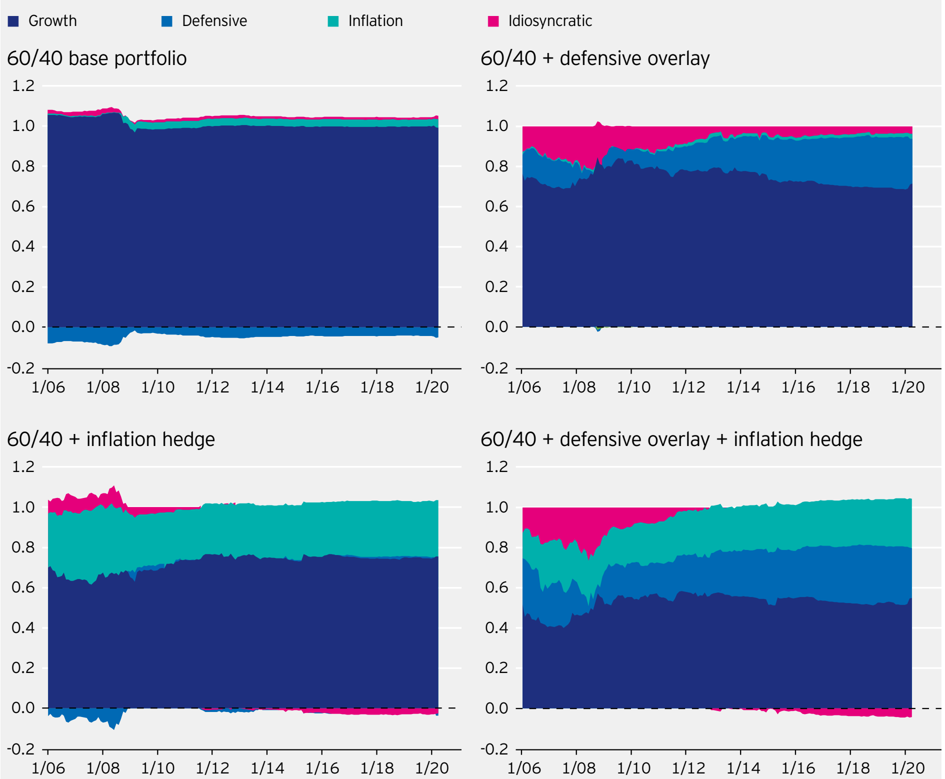
The chart decomposes the volatility of the macro factor-mimicking portfolio for defensive (left) and inflation (right) by macroeconomic factor contributions.
Sources: Invesco, Bloomberg, Goldman Sachs. Sample period 31 January 2006 to 31 May 2020.

Adding a defensive overlay
To keep the analysis simple, we add the defensive MFMP to the 60/40 stock-bond allocation using the exact defensive MFMP weights given in table 1. The addition of the defensive MFMP comes with a noticeable reduction in growth risk, as we can infer from the macro factor risk decomposition in figure 4 (upper right). Moreover, this addition has a considerable impact on the ensuing portfolio's risk-return profile. In the absence of the defensive overlay, the 60/40 portfolio operates at a 10% volatility level (see table 2). Given a Sharpe ratio of 0.54, it delivered some 6.2% annualized return over the sample period. Adding the defensive overlay barely affects the volatility level (which is slightly down to 9.44%) but crucially mitigates tail risk; maximum drawdown is considerably cut (-25.58%), which represents a reduction of more than 10 percentage points relative to that of the 60/40 portfolio (-36.54%). As a result, the annualized return is almost twice as high as that of the 60/40 (11.42% versus 6.22%).

Adding a diversified inflation hedge overlay
The effect of adding an inflation hedge clearly shows in the macro factor risk decomposition (figure 5, lower left). However, the inflation hedge portfolio is not a source of extra return in the backtest period. The combination with the 60/40 slightly raises volatility and tail risk due to the consideration of commodity assets. Therefore, we also provide performance statistics of a strategy variant that scales the full allocation such that the ensuing volatility is comparable to the one of the 60/40 base allocation. The annualized return of the scaled strategy is 3.97%, which is 225bp below the 60/40 portfolio's return. Still, the drawdown is likewise severe and comes in at -38.15%. Obviously, one would need to consult a longer history to better gauge the actual benefit of inflation hedging, as the considered sample period is lacking pronounced inflationary regimes.

Diversifying growth risk through combining defensive and inflation hedge overlays
Given the difficulties in predicting the economic environment, diversifying macroeconomic factor risk seems a natural path to follow. We thus investigate adding both overlays, defensive and inflation hedge, to the 60/40 allocation. First, we observe a fairly balanced risk profile through time, where growth risk ceases to be dominant most of the time. From a performance perspective, there is a slight increase in volatility risk but still a considerable decrease in tail

Figure 5
Diversifying a traditional asset allocation with macro factor overlays



The figure decomposes the volatility of a 60/40 stock-bond allocation by macroeconomic factor contributions (upper left). The remaining charts refer to 60/40 stock-bond allocations amended by a defensive overlay (upper right), an inflation hedge (lower left), as well as a combination of the two (lower right).
Sources: Invesco, Bloomberg, Goldman Sachs. Sample period 31 January 2006 to 31 May 2020.

Table 2
Performance statistics for macro factor-based allocations

	Defensive MFMP	Inflation hedge	60/40	60/40 + Defensive	60/40 + Inflation Hedge	60/40 + Defensive + Inflation Hedge	60/40 + Defensive (Scaled)	60/40 + Inflation Hedge (Scaled)	60/40 + Defensive + Inflation Hedge (Scaled)
Net return p.a.	5.92%	0.59%	6.22%	11.42%	5.48%	10.92%	11.91%	3.97%	8.96%
Volatility p.a.	4.41%	3.93%	10.04%	9.44%	12.18%	11.56%	10.33%	10.36%	10.14%
Sharpe Ratio	1.07	-0.12	0.54	1.08	0.41	0.86	1.04	0.32	0.79
Max Drawdown	-6.19%	-12.75%	-36.54%	-25.28%	-39.35%	-30.45%	-28.77%	-38.15%	-30.66%
Calmar Ratio	0.96	0.09	0.17	0.45	0.14	0.36	0.41	0.1	0.29

The table provides simulated performance figures for macro factor-based multi-asset multi-factor strategies from the perspective of a US-dollar investor. This model does not factor in all economic and market conditions that can impact results.
Source: Invesco, Bloomberg, Goldman Sachs. Period from January 2006 to May 2020. **The figures refer to simulated past performance and past performance is not a reliable indicator of future performance.**

Table 3
Macro factor allocations – performance by regime

	Rising Growth + Rising Inflation	Rising Growth + Falling Inflation	Falling Growth + Rising Inflation	Falling Growth + Falling Inflation
Growth	10.9%	4.8%	-2.8%	-17.8%
Defensive Overlay	1.2%	5.6%	8.8%	14.1%
Inflation Hedge	4.8%	-2.3%	4.7%	-8.0%
60/40	26.9%	17.5%	-18.9%	-29.5%
60/40 + Defensive Overlay	27.1%	21.4%	-11.5%	-16.3%
60/40 + Inflation Hedge	30.7%	13.6%	-15.5%	-38.3%
60/40 + Defensive + Inflation	31.2%	17.7%	-7.8%	-25.0%
60/40 + Defensive Overlay (scaled)	29.0%	23.5%	-13.2%	-18.4%
60/40 + Inflation Hedge (scaled)	24.7%	11.8%	-14.0%	-32.6%
60/40 + Defensive + Inflation (scaled)	26.5%	15.8%	-7.4%	-22.5%

The table displays annualized excess returns of several macro factor allocations performance under four different growth-inflation regimes.
Source: Invesco, Bloomberg, Goldman Sachs. Sample period from January 2006 to May 2020. **The figures refer to simulated past performance and past performance is not a reliable indicator of future performance.**

risk. The annualized return is almost 50% higher than in the 60/40 base case, resulting in a Sharpe ratio of 0.79. Also, the drawdown is reduced by some 6 percentage points to -30.66%.

Lastly, we investigate how the different macro factor strategies perform in the four growth-inflation regimes defined earlier; see table 3. Given its concentration in growth risk, we find the 60/40 portfolio outperforming in rising growth environments and underperforming when growth falls. As expected, the defensive macro factor-mimicking portfolio is particularly beneficial in both negative growth regimes. In fact, the regime-specific performance analysis highlights that the 60/40 with defensive overlay is on par with the pure 60/40 in rising growth environments, but considerably better in the two falling growth periods.

As for the efficacy of the inflation hedge, we observe that the corresponding macro factor-mimicking portfolio indeed earns positive returns in inflationary regimes and negative returns in deflationary regimes. Adding this inflation hedge to the 60/40 allocation, we note that the rising growth/falling inflation regime sees similar returns for this and the base portfolio. Yet, under the falling growth/rising inflation regime, we observe a return benefit for the inflation-hedged strategy. Judging by the scaled 60/40 + inflation hedge, the regime return is half that of the pure 60/40 portfolio (-14.0% vs. -18.9%). Obviously, if one is expecting a falling growth/rising inflation period, enhancing the 60/40 allocation through the addition of defensive and inflation hedge overlays is the method of choice, as demonstrated by a historic regime-specific return of -7.4% (based on the scaled version).

Building out macro factor-mimicking portfolios from a straightforward growth-inflation regime perspective enables effective diversification.

Conclusion

Style factors are often considered meaningful diversifiers and can help achieve various investor objectives; see Dichtl, Drobetz, Lohre and Rother (2021). In this article, we have looked at such investments through the overarching lens of macro factors that ultimately govern the dynamics of asset class and style factor returns. Building out macro factor-mimicking portfolios from a straightforward growth-inflation regime perspective enables effective diversification of traditional stock-bond allocations versus growth and inflation risks.

References

Amato, L., and H. Lohre (2020): Diversifying macroeconomic factors – for better or for worse, Working paper.

Bernardi, S., M. Leippold and H. Lohre (2018): Maximum diversification strategies along commodity risk factors, European Financial Management, 24(1), 53-78.

Chen, N., R. Roll and S.A. Ross (1986): Economic forces and the stock market, Journal of Business, 59(3), 383-403.

Dichtl, H., W. Drobetz, H. Lohre and C. Rother (2021): Active factor completion strategies, Journal of Portfolio Management, forthcoming.

Greenberg, D., A. Babu and A. Ang (2016): Factors to assets: Mapping factor exposures to asset allocations, Journal of Portfolio Management, Special Issue, 18-27.

Kothe, J., H. Lohre, and C. Rother (2021): Rates factors and global asset allocation, Journal of Fixed Income, forthcoming.

Notes

1 See Amato and Lohre (2020) for a comparison of associated diversified macro factor risk parity strategies.
2 Performing statistical clustering based on the multi-asset multi-factor dataset underlying this study, we find support for the relevance of the three factors growth, inflation and defensive; results available upon request.
3 Note that the rising growth / rising inflation regime prevails in about 40% of our sample period while the remaining three regimes each prevail in some 20% of the months. Hence, the derived statistics are not based on overly thin data.
4 By construction, the MSCI ACWI comes with inflated growth exposures and is thus not a suitable anchor.

About the authors

Dr. Harald Lohre
Director of Research
Invesco Quantitative Strategies
Harald Lohre and his team are responsible for maintaining and evolving the quantitative models that drive the investment decisions within multi-factor equity and balanced investment products.

Scott Hixon, CFA®
Portfolio Manager and Head of Research
Invesco Global Asset Allocation
Scott Hixon is a Portfolio Manager and the Head of Research for Invesco's Global Asset Allocation team, which invests in stock, bond and commodity markets worldwide. Mr. Hixon oversees and helps steer the team's research initiatives in the areas of model and strategy development, as well as portfolio construction.

Jay Raol, Ph.D.
Head of Fixed Income Factors
Systematic & Factor Investment Group
Jay Raol and his team researches and manages systematic and factor based strategies in global fixed income and currency markets.

Alexander Swade
Ph.D. Candidate, Lancaster University, and
Invesco Quantitative Strategies
In a joint research initiative between Lancaster University and Invesco Quantitative Strategies, Alexander Swade investigates factor-based investment strategies across different asset classes and within equities.

Hua Tao, Ph.D., CFA®
Research Analyst
Invesco Global Asset Allocation
Hua Tao is a Research Analyst for the Invesco Global Asset Allocation team, responsible for model research and strategy development.

Scott Wolle, CFA®
Head of Systematic & Factor Investing
Invesco
Scott Wolle serves as the Head of Systematic & Factor Investing, which includes equity, fixed income, and macro strategies along with custom indices.

Data appendix

Name	Description	Ticker	Source	Construction details
Equities				
ACWI	MSCI ACWI Net TR Local index	NDLEACWF	Bloomberg	
US-ACWI	MSCI USA TR USD Index minus MSCI ACWI Net TR Local Index	NDDLUS, NDLEACWF	Bloomberg	NDDLUS - NDLEACWF
EAFE-ACWI	MSCI EAFE TR LCL Index minus MSCI ACWI Net TR Local Index	NDDLEAFE, NDLEACWF	Bloomberg	NDDLEAFE - NDLEACWF
EM-ACWI	MSCI EM TR LCL Index minus MSCI ACWI Net TR Local Index	NDLEEGF, NDLEACWF	Bloomberg	NDLEEGF - NDLEACWF
Cyclicals-Defensives	ACWI CYCLICAL SECTORS- ACWI DEFENSIVE SECTORS	MXCXDRN	Bloomberg	
Quality	MSCI ACWI Quality USD minus MSCI ACWI Net TR USD Index	M1WDQU, NDUEACWF	Bloomberg	M1WDQU - vol.adj. NDUEACWF
Momentum	MSCI ACWI Momentum USD minus MSCI ACWI Net TR USD Index	M1WD000$, NDUEACWF	Bloomberg	M1WD000$ - vol.adj. NDUEACWF
Value	MSCI ACWI Value USD minus MSCI ACWI Net TR USD Index	M1WD000V, NDUEACWF	Bloomberg	M1WD000V - vol.adj. NDUEACWF
Low Volatility	MSCI ACWI Minimum Volatility USD minus MSCI ACWI Net TR USD Index	M00IWD$O, NDUEACWF	Bloomberg	M00IWD$O - vol.adj. NDUEACWF
Fixed Income				
US 10Y Tsy	Bloomberg Barclays US Treasury Bellwethers 10 Year TR Index Value Unhedged USD	BW10TRUU	Bloomberg	
Cash	USD 3 Month T-Bill	USGG3M	Bloomberg	
TIPS	US TIPS TR	I01551US	Bloomberg	
IG Credit	Bloomberg Barclays US Agg Corp excess return	LUACER	Bloomberg	
HY Credit	Bloomberg Barclays US Corporat HY excess return	LF98ER	Bloomberg	
EM Credit	J.P. Morgan EMBI Global TR minus US Treasury	JPEIGLBL, LUATTRUU	Bloomberg	JPEIGLBL - vol.adj. LUATTRUU
Rates Value	Goldman Sachs Rates Value Strategy	GSIRVA03	GS	
Rates Momentum	Goldman Sachs Rates Momentum Strategy	GSIRTR03	GS	
Rates Quality	Goldman Sachs Rates Quality Strategy	GS Interest Rates Curve C0210	GS	
Rates Carry	Goldman Sachs Rates Carry Strategy	GSIRCA03	GS	
Commodities				
Precious Metals	Bloomberg Precious Metals Subindex	BCOMPR	Bloomberg	
Industrial Metals	Bloomberg Industrial Metals Subindex	BCOMIN	Bloomberg	
Energy	Bloomberg Energy Subindex	BCOMEN	Bloomberg	
Agriculture	Bloomberg Agriculture Subindex	BCOMAG	Bloomberg	
Carry	Goldman Sachs Commodity Carry Strategy	GS Macro Carry Index RP14	GS	
Quality	Bloomberg Roll Select Commodity Index minus Bloomberg Commodity Index	BCOMRST, BCOMTR	Bloomberg	BCOMRST - BCOMTR
Momentum	Goldman Sachs Commodity Momentum Strategy	GS Macro Momentum Index RP15	GS	
Value	Goldman Sachs Commodity Value Strategy	GS Commodity COT Strategy COT3	GS	
Currencies				
Developed Markets	MSCI EAFE Currency USD Index	MXEA0CX0	Bloomberg	
Emerging Markets	MSCI Emerging Markets Currency USD Index	MXEF0CX0	Bloomberg	
Carry	Goldman Sachs FX Carry Strategy	GS FX Carry C0115	GS	
Value	Goldman Sachs FX Value Strategy	GS FX Value C0114	GS	
Momentum	Goldman Sachs FX Momentum Strategy	GS FX Trend C0038	GS	
Macro factors				
Growth	MSCI ACWI Net TR Local index	NDLEACWF	Bloomberg	
Defensive	Bloomberg Barclays US Treasury Bellwethers 10 Year TR Index Value Unhedged USD	BW10TRUU	Bloomberg	
Inflation	USTreasuryTIP minus BBUSTreasury	SPBDUP3T, LT01TRUU	Bloomberg	SPBDUP3T - LT01TRUU

References

Amato, Livia and Harald Lohre (2020). "Diversifying macroeconomic factors—for better or for worse." *Available at SSRN: https://ssrn.com/abstract=3730154*.

Ang, Andrew and Geert Bekaert (2004). "How regimes affect asset allocation." *Financial Analysts Journal* 60 (2), 86–99.

Ang, Andrew and Joseph Chen (2002). "Asymmetric correlations of equity portfolios." *Journal of Financial Economics* 63 (3), 443–494.

Arnott, Robert D., Jason Hsu, and Philip Moore (2005). "Fundamental indexation." *Financial Analysts Journal* 61 (2), 83–99.

Asness, Clifford S., Andrea Frazzini, Ronen Israel, Tobias J. Moskowitz, and Lasse H. Pedersen (2018). "Size matters, if you control your junk." *Journal of Financial Economics* 129 (3), 479–509.

Asness, Clifford S., Andrea Frazzini, and Lasse H. Pedersen (2019). "Quality minus junk." *Review of Accounting Studies* 24 (1), 34–112.

Baltussen, Guido, Laurens Swinkels, and Pim Van Vliet (2021). "Global factor premiums." *Journal of Financial Economics* 142 (3), 1128–1154.

Banz, Rolf W. (1981). "The relationship between return and market value of common stocks." *Journal of Financial Economics* 9 (1), 3–18.

Barillas, Francisco, Raymond Kan, Cesare Robotti, and Jay Shanken (2020). "Model comparison with sharpe ratios." *Journal of Financial and Quantitative Analysis* 55 (6), 1840–1874.

Barillas, Francisco and Jay Shanken (2017). "Which alpha?." *Review of Financial Studies* 30 (4), 1316–1338.

Bartram, Söhnke M., Harald Lohre, Peter F. Pope, and Ananthalakshmi Ranganathan (2021). "Navigating the factor zoo around the world: an institutional investor perspective." *Journal of Business Economics* 91 (5), 655–703.

Bass, Robert, Scott Gladstone, and Andrew Ang (2017). "Total portfolio factor, not just asset, allocation." *Journal of Portfolio Management* 43 (5), 38–53.

Basu, Sanjoy (1977). "Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis." *Journal of Finance* 32 (3), 663–682.

Bernardi, Simone, Markus Leippold, and Harald Lohre (2018). "Maximum diversification strategies along commodity risk factors." *European Financial Management* 24 (1), 53–78.

Bessembinder, Hendrik, Aaron Burt, and Christopher M. Hrdlicka (2021). "Time series variation in the factor zoo." *Available at SSRN: https://ssrn.com/abstract=3992041*.

— (2022). "Factor returns and out-of-sample alphas: Factor construction matters." *Available at SSRN: https://ssrn.com/abstract=4281769*.

Black, Fischer and Robert B. Litterman (1991). "Asset Allocation: Combining Investor Views with Market Equilibrium." *Journal of Fixed Income* 1 (2), 7–18.

— (1992). "Global portfolio optimization." *Financial Analysts Journal* 48 (5), 28–43.

Blin, Olivier, Florian Ielpo, Joan Lee, and Jerome Teiletche (2021). "Alternative risk premia timing: A point-in-time macro, sentiment, valuation analysis." *Journal of Systematic Investing* 1 (1), 52–72.

Blitz, David (2023). "The Cross-Section of Factor Returns." *Journal of Portfolio Management, forthcoming*.

Blitz, David and Matthias X. Hanauer (2020). "Settling the size matter." *The Journal of Portfolio Management* 47 (2), 99–112.

Blitz, David, Matthias X. Hanauer, Iman Honarvar, Rob Huisman, and Pim van Vliet (2023). "Beyond Fama-French factors: Alpha from short-term signals." *Financial Analysts Journal*, forthcoming.

Blitz, David, Matthias X. Hanauer, and Milan Vidojevic (2020). "The idiosyncratic momentum anomaly." *International Review of Economics & Finance* 69, 932–957.

Blitz, David, Joop Huij, and Martin Martens (2011). "Residual momentum." *Journal of Empirical Finance* 18 (3), 506–521.

Bouchey, Paul, Vassilii Nemtchinov, Alex Paulsen, and David M. Stein (2012). "Volatility harvesting: Why does diversifying and rebalancing create portfolio growth?." *The Journal of Wealth Management* 15 (2), 26–35.

Carhart, Mark (1997). "On persistence in mutual fund performance." *Journal of Finance* 52 (1), 57–82.

Chen, Andrew Y. and Tom Zimmermann (2022). "Open Source Cross-Sectional Asset Pricing." *Critical Finance Review* 11 (2), 207–264.

Chen, Nai-Fu, Richard Roll, and Stephen A. Ross (1986). "Economic forces and the stock market." *Journal of Business* 59 (3), 383–403.

Chordia, Tarun and Lakshmanan Shivakumar (2002). "Momentum, business cycle, and time-varying expected returns." *Journal of Finance* 57 (2), 985–1019.

Chousakos, Kyriakos and Daniel Giamouridis (2020). "Harvesting macroeconomic risk premia." *Journal of Portfolio Management* 46 (6), 93–109.

Cochrane, John H. (2011). "Presidential address: Discount rates." *Journal of Finance* 66 (4), 1047–1108.

Cooper, Ilan, Andreea Mitrache, and Richard Priestley (2022). "A global macroeconomic risk model for value, momentum, and other asset classes." *Journal of Financial and Quantitative Analysis* 57 (1), 1–30.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill (2008). "Asset growth and the cross-section of stock returns." *Journal of Finance* 63 (4), 1609–1651.

D'avolio, Gene (2002). "The market for borrowing stock." *Journal of Financial Economics* 66 (2-3), 271–306.

De Wit, Dirk P. M. (1998). "Naive diversification." *Financial Analysts Journal* 54 (4), 95–100.

Deguest, Romain, Lionel Martellini, and Attilio Meucci (2013). "Risk parity and beyond – from asset allocation to risk allocation decisions." *EDHEC-Risk Institute (June).*

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal (2009). "Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?." *The Review of Financial Studies* 22 (5), 1915–1953.

Dichtl, Hubert, Wolfgang Drobetz, Harald Lohre, and Carsten Rother (2021). "Active factor completion strategies." *Journal of Portfolio Management* 47 (2), 9–37.

Erb, Claude B. and Campbell R. Harvey (2006). "The strategic and tactical value of commodity futures." *Financial Analysts Journal* 62 (2), 69–97.

Fama, Eugene F. (1998). "Determining the number of priced state variables in the ICAPM." *Journal of Financial and Quantitative Analysis* 33 (2), 217–231.

Fama, Eugene F. and Kenneth R. French (1989). "Business conditions and expected returns on stocks and bonds." *Journal of Financial Economics* 25 (1), 23–49.

— (1993). "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics* 33 (1), 3–56.

— (1996). "Multifactor explanations of asset pricing anomalies." *Journal of Finance* 51 (1), 55–84.

— (2012). "Size, value, and momentum in international stock returns." *Journal of Financial Economics* 105 (3), 457–472.

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

— (2016). “Dissecting anomalies with a five-factor model.” *Review of Financial Studies* 29 (1), 69–103.

— (2018). “Choosing factors.” *Journal of Financial Economics* 128 (2), 234–252.

Fama, Eugene F. and James D. MacBeth (1973). “Risk, return, and equilibrium: Empirical tests.” *Journal of Political Economy* 81 (3), 607–636.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu (2020). “Taming the factor zoo: A test of new factors.” *Journal of Finance* 75 (3), 1327–1370.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken (1989). “A test of the efficiency of a given portfolio.” *Econometrica: Journal of the Econometric Society*, 1121–1152.

Green, Jeremiah, John R. M. Hand, and X. Frank Zhang (2017). “The characteristics that provide independent information about average US monthly stock returns.” *Review of Financial Studies* 30 (12), 4389–4436.

Griffin, John M. (2002). “Are the Fama and French factors global or country specific?.” *Review of Financial Studies* 15 (3), 783–803.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu (2020). “Empirical asset pricing via machine learning.” *Review of Financial Studies* 33 (5), 2223–2273.

Gupta, Tarun and Bryan Kelly (2019). “Factor momentum everywhere.” *Journal of Portfolio Management* 45 (3), 13–36.

Hallerbach, Winfried G. (2014). “Disentangling rebalancing return.” *Journal of Asset Management* 15 (5), 301–316.

Hanauer, Matthias X. (2020). “A comparison of global factor models.” *Available at SSRN: https://ssrn.com/abstract=3546295.*

Hanauer, Matthias X. and Tobias Kalsbach (2023). “Machine learning and the cross-section of emerging market stock returns.” *Emerging Markets Review* 55, 101022.

Hanauer, Matthias X. and Martin Linhart (2015). “Size, value, and momentum in emerging market stock returns: Integrated or segmented pricing?.” *Asia-Pacific Journal of Financial Studies* 44 (2), 175–214.

Harvey, Campbell R., Yan Liu, and Heqing Zhu (2016). “. . . and the cross-section of expected returns.” *Review of Financial Studies* 29 (1), 5–68.

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

Hou, Kewei, Chen Xue, and Lu Zhang (2015). “Digesting anomalies: An investment approach.” *Review of Financial Studies* 28 (3), 650–705.

— (2020). “Replicating anomalies.” *Review of Financial Studies* 33 (5), 2019–2133.

Huberman, Gur, Shmuel Kandel, and Robert F. Stambaugh (1987). "Mimicking portfolios and exact arbitrage pricing." *Journal of Finance* 42 (1), 1–9.

Ilmanen, Antti, Ronen Israel, Tobias J. Moskowitz, Ashwin K. Thapar, and Rachel Lee (2021). "How do factor premia vary over time? A century of evidence." *Journal of Investment Management* 19 (4), 15–57.

Ilmanen, Antti, Thomas Maloney, and Adrienne Ross (2014). "Exploring macroeconomic sensitivities: How investments respond to different economic environments." *Journal of Portfolio Management* 40 (3), 87–99.

Jacobs, Heiko and Sebastian Müller (2018). "... And Nothing Else Matters? On the Dimensionality and Predictability of International Stock Returns." *Available at SSRN: https://ssrn.com/abstract=2845306.*

Jegadeesh, Narasimhan and Sheridan Titman (1993). "Returns to buying winners and selling losers: Implications for stock market efficiency." *Journal of Finance* 48 (1), 65–91.

Jensen, Theis Ingerslev, Bryan T. Kelly, and Lasse Heje Pedersen (2023). "Is there a replication crisis in finance?." *Journal of Finance, Forthcoming.*

Jurczenko, Emmanuel and Jérôme Teiletche (2018). "Active risk-based investing." *Journal of Portfolio Management* 44 (3), 56–65.

— (2022). "Macro factor mimicking portfolios." *Available at SSRN 3363598.*

Kaiser, Lars and Georg Peter (2022). "Riding the 1/N Premium." *The Journal of Investing* 31 (2), 94–109.

Kaplan, Paul D. (2008). "Why fundamental indexation might—or might not—work." *Financial Analysts Journal* 64 (1), 32–39.

Keim, Donald B (1983). "Size-related anomalies and stock return seasonality: Further empirical evidence." *Journal of Financial Economics* 12 (1), 13–32.

Kirby, Chris (2019). "The value premium and expected business conditions." *Finance Research Letters* 30, 360–366.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh (2018). "Interpreting factor models." *Journal of Finance* 73 (3), 1183–1223.

Kritzman, Mark, Sebastien Page, and David Turkington (2012). "Regime shifts: Implications for dynamic strategies (corrected)." *Financial Analysts Journal* 68 (3), 22–39.

Lamont, Owen A. (2001). "Economic tracking portfolios." *Journal of Econometrics* 105 (1), 161–184.

Lettau, Martin and Markus Pelger (2020). "Factors that fit the time series and cross-section of stock returns." *Review of Financial Studies* 33 (5), 2274–2325.

Lintner, John (1965). "Security prices, risk, and maximal gains from diversification." *Journal of Finance* 20 (4), 587–615.

Lohre, Harald, Robert S Hixon, Jay H Raol, Alexander Swade, Hua Tao, and Scott Wolle (2020). "Investing through a macro factor lens." *Risk & Reward*, 26–34.

Lohre, Harald, Heiko Opfer, and Gábor Ország (2014). "Diversifying risk parity." *Journal of Risk* 16 (5), 53–79.

Lohre, Harald, Carsten Rother, and Kilian A. Schäfer (2020). "Hierarchical risk parity: Accounting for tail dependencies in multi-asset multi-factor allocations." *Machine Learning for Asset Management: New Developments and Financial Applications*. Ed. by Emmanuel Jurczenko. John Wiley & Sons, Ltd. Chap. 9, 329–368.

Longin, Francois and Bruno Solnik (2001). "Extreme correlation of international equity markets." *Journal of Finance* 56 (2), 649–676.

Maeso, Jean-Michel and Lionel Martellini (2020). "Measuring portfolio rebalancing benefits in equity markets." *The Journal of Portfolio Management* 46 (4), 94–109.

Maillard, Sébastien, Thierry Roncalli, and Jérôme Teiletche (2010). "The properties of equally weighted risk contribution portfolios." *Journal of Portfolio Management* 36 (4), 60–70.

Malladi, Rama and Frank J. Fabozzi (2017). "Equal-weighted strategy: Why it outperforms value-weighted strategies? Theory and evidence." *Journal of Asset Management* 18 (3), 188–208.

Markowitz, Harry (1952). "Portfolio selection." *Journal of Finance* 7 (1), 77–91.

McLean, R. David and Jeffrey Pontiff (2016). "Does academic research destroy stock return predictability?." *Journal of Finance* 71 (1), 5–32.

Meucci, Attilio (2009). "Managing diversification." *Risk* 22, 74–79.

— (2010). "Black–Litterman approach." *Encyclopedia of Quantitative Finance*.

Meucci, Attilio, Alberto Santangelo, and Romain Deguest (2015). "Risk budgeting and diversification based on optimized uncorrelated factors." *Risk* 11, 70–75.

Moskowitz, Tobias J, Yao Hua Ooi, and Lasse Heje Pedersen (2012). "Time series momentum." *Journal of Financial Economics* 104 (2), 228–250.

Mossin, Jan (1966). "Equilibrium in a capital asset market." *Econometrica* 34 (4), 768–783.

Murtagh, Fionn and Pierre Legendre (2014). "Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion?." *Journal of Classification* 31, 274–295.

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

Pae, Yuntaek and Navid Sabbaghi (2015). "Equally weighted portfolios vs. value weighted portfolios: Reasons for differing betas." *Journal of Financial Stability* 18, 203–207.

Partovi, M. Hossein and Michael Caputo (2004). "Principal portfolios: Recasting the efficient frontier." *Economics Bulletin* 7 (3), 1–10.

Patton, Andrew J. and Allan Timmermann (2010). "Monotonicity in asset returns: New tests with applications to the term structure, the CAPM, and portfolio sorts." *Journal of Financial Economics* 98 (3), 605–625.

Penrose, Roger (1955). "A generalized inverse for matrices." *Mathematical Proceedings of the Cambridge Philosophical Society* 51 (3), 406–413.

Perold, André F. (2007). "Fundamentally flawed indexing." *Financial Analysts Journal* 63 (6), 31–37.

Perold, Andre F. and William F. Sharpe (1988). "Dynamic strategies for asset allocation." *Financial Analysts Journal* 44 (1), 16–27.

Plyakha, Yuliya, Raman Uppal, and Grigory Vilkov (2021). "Equal or value weighting? Implications for asset-pricing tests." *Financial Risk Management and Modeling*. Ed. by Constantin Zopounidis, Ramzi Benkraiem, and Iordanis Kalaitzoglou. Springer, 295–347.

Polk, Christopher, Mo Haghbin, and Alessio De Longis (2020). "Time-series variation in factor premia: The influence of the business cycle." *Journal of Investment Management* 18 (1), 69–89.

Pontiff, Jeffrey and Lawrence D. Schall (1998). "Book-to-market ratios as predictors of market returns." *Journal of Financial Economics* 49 (2), 141–160.

Pukthuanthong, Kuntara, Richard Roll, Junbo L. Wang, and Tengfei Zhang (2019). "A New Method for Factor-Mimicking Portfolio Construction." *Available at SSRN: https://ssrn.com/abstract=3*

Qian, Edward E. (2006). "On the financial interpretation of risk contribution: Risk budgets do add up." *Journal of Investment Management* 4 (4), 41–51.

Rapach, David E., Jack K. Strauss, and Guofu Zhou (2013). "International stock return predictability: What is the role of the United States?." *Journal of Finance* 68 (4), 1633–1662.

Reinganum, Marc R. (1981). "Misspecification of capital asset pricing: Empirical anomalies based on earnings' yields and market values." *Journal of Financial Economics* 9 (1), 19–46.

Roll, Richard (1983). "Vas ist das?." *The Journal of Portfolio Management* 9 (2), 18–28.

Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein (1985). "Persuasive evidence of market inefficiency." *Journal of Portfolio Management* 11 (3), 9–16.

Ross, Stephen A. (1976). "The arbitrage theory of capital asset pricing." *Journal of Economic Theory* 13 (3), 341–360.

Scherer, Bernd and Matthias Apel (2020). "Business cycle–related timing of alternative risk premia strategies." *Journal of Alternative Investments* 22 (4), 8–24.

Sharpe, William F. (1964). "Capital asset prices: A theory of market equilibrium under conditions of risk." *Journal of Finance* 19 (3), 425–442.

Siegel, Jeremy (2006). "The noisy market hypothesis." *Wall Street Journal* 14, A14.

Soebhag, Amar, Bart Van Vliet, and Patrick Verwijmeren (2023). "Non-Standard Errors in Asset Pricing: Mind Your Sorts." *Available at SSRN: https://ssrn.com/abstract=4136672*.

Stambaugh, Robert F. and Yu Yuan (2017). "Mispricing factors." *Review of Financial Studies* 30 (4), 1270–1315.

Swade, Alexander, Harald Lohre, Matthias X. Hanauer, and David Blitz (2003). "Factor zoo (.zip)." *working paper*.

Swade, Alexander, Harald Lohre, Mark Shackleton, Sandra Nolte, Scott Hixon, and Jay Raol (2021). "Macro factor investing with style." *Journal of Portfolio Management* 48 (2), 80–104.

Swade, Alexander, Sandra Nolte, Mark Shackleton, and Harald Lohre (2023). "Why do equally weighted portfolios beat value-weighted ones?." *Journal of Portfolio Management* 49 (5), 167–187.

Tobek, Ondrej and Martin Hronec (2021). "Does it pay to follow anomalies research? Machine learning approach with international evidence." *Journal of Financial Markets* 56, 100588.

Treynor, Jack L. (1961). "Market value, time, and risk." *Available at SSRN: https://ssrn.com/abstract=2600356*.

Van Vliet, Pim and David Blitz (2011). "Dynamic strategic asset allocation: Risk and return across the business cycle." *Journal of Asset Management* 12 (5), 360–375.

Van Vliet, Pim and Jan De Koning (2017). "High returns from low risk: A remarkable stock market paradox." John Wiley & Sons.

Ward, Joe H., Jr. (1963). "Hierarchical grouping to optimize an objective function." *Journal of the American Statistical Association* 58 (301), 236–244.

Willenbrock, Scott (2011). "Diversification return, portfolio rebalancing, and the commodity return puzzle." *Financial Analysts Journal* 67 (4), 42–49.