

Dichtl, Hubert; Drobetz, Wolfgang; Wendt, Viktoria-Sophie

Article — Published Version

How to build a factor portfolio: Does the allocation strategy matter?

European Financial Management

Provided in Cooperation with:

John Wiley & Sons

Suggested Citation: Dichtl, Hubert; Drobetz, Wolfgang; Wendt, Viktoria-Sophie (2021) : How to build a factor portfolio: Does the allocation strategy matter?, European Financial Management, ISSN 1468-036X, Wiley, Hoboken, NJ, Vol. 27, Iss. 1, pp. 20-58, <https://doi.org/10.1111/eufm.12264>

This Version is available at:

<https://hdl.handle.net/10419/230157>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>

How to build a factor portfolio: Does the allocation strategy matter?

Hubert Dichtl¹ | Wolfgang Drobetz¹ | Viktoria-Sophie Wendt²

¹Faculty of Business, University of Hamburg, Hamburg, Germany

²BlackRock Investment Management Limited, London, UK

Correspondence

Wolfgang Drobetz, Faculty of Business, University of Hamburg, Moorweidenstr, 18, 20148 Hamburg, Germany.

Email: wolfgang.drobetz@uni-hamburg.de

Abstract

Factor-based allocation embraces the idea of factors, as opposed to asset classes, as the ultimate building blocks of investment portfolios. We examine whether there is a superior way of combining factors in a portfolio and provide a comparison of factor-based allocation strategies within a multiple testing framework. Factor-based allocation is profitable beyond exploiting genuine risk premia, even when applying multiple testing corrections. Investment portfolios can be efficiently diversified using factor-based allocation strategies, as demonstrated by robust economic performance over various economic scenarios. The naïve equally weighted factor portfolio, albeit simple and cost-efficient, cannot be outperformed by more sophisticated allocation strategies.

KEYWORDS

factor-based allocation, multiple testing

JEL CLASSIFICATION

G11; G23

We thank two anonymous referees, Geert Bekaert, John Doukas (editor), Tizian Otto, Tatjana Puhon, and Henning Schröder for helpful comments.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *European Financial Management* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Building on the seminal work of Markowitz (1952), investors have traditionally focused on diversifying across broad asset classes, such as equities and bonds, when building their investment portfolios in order to balance risks and rewards. The premise underlying the asset allocation decision is that, while efficiently diversifying away unrewarded (idiosyncratic) risks, the asset classes considered are still subject to an inherent return premium that can be explained by some underlying common factors. The capital asset pricing model (Lintner, 1965; Mossin, 1966; Sharpe, 1964) promoted the “market” as the only explanatory factor. However, ever since the introduction of the arbitrage pricing theory (Ross, 1976), assets have been thought of as a bundle of multiple factors that reflect different risks and rewards (Ang, 2014; Ang, Goetzmann, & Schaefer, 2009; Ferson & Harvey, 1993).

Koedijk, Slager, and Stork (2016a) contrast three approaches to integrating the concept of factor bundles into the investment management process. The first, so-called risk due diligence approach, simply uses the factor methodology to gain a better understanding of the factor exposures within a previously determined investment portfolio, and to potentially revise the asset allocation accordingly. The second approach, factor tilting, is currently the most widely used approach and refers to actively tweaking a portfolio's exposure toward certain factors based on the existing asset allocation by either introducing factors biases within the assets or complementing the assets with “pure” factors (Amenc, Deguest, Goltz, Lodh, & Martellini, 2014; Dichtl, Drobetz, Lohre, & Rother, 2020). Lastly, factor optimization uses the factors, instead of asset classes, as the ultimate building blocks of investment portfolios, thus deciding on a factor-based as opposed to asset-based allocation. Scarred by the last financial crisis, when presumably “well-diversified” asset-based portfolios declined substantially, factor-based allocation is becoming increasingly popular with institutional investors seeking to effectively diversify their portfolios (Koedijk, Slager, & Stork, 2016b; Ung & Kang, 2015).

While the advantages and disadvantages of factor-based allocation are fiercely debated (Idzorek & Kowara, 2013; Ilmanen & Kizer, 2012), the question how to optimally combine factors in a portfolio has remained largely unanswered. Some recent studies use a heuristic approach such as a portfolio of equally weighted or equal-volatility weighted factors (Bender, Briand, Nielsen, & Stefek, 2010; Ilmanen & Kizer, 2012). Other academic studies revert to risk-minimization strategies such as the global minimum variance portfolio (Amenc et al., 2014), or analyze efficient frontiers from mean-variance optimizations (Brière & Szafarz, 2017; Clarke, de Silva, & Murdock, 2005; Idzorek & Kowara, 2013). While each approach has its pros and cons, a thorough comparison of these allocation strategies against the backdrop of factor-based portfolio formation is pending.

To the best of our knowledge, our study is the first to jointly compare the performance of a comprehensive set of tradable factor-based allocation strategies in order to address two research questions. First, is factor-based allocation able to provide superior portfolio returns? And second, does the specific allocation strategy applied matter much, that is, is there a superior way of combining factors in a portfolio?¹ Our analysis assumes an investor who allocates to a comprehensive set of global equity and fixed income factors and applies a variety of factor optimization strategies that are either estimation-free or require the estimation of some factor moments. In addition, we use Hansen's (2005) test for superior predictive ability (SPA test) and its extensions to avoid a common bias in statistical inference referred to as “data snooping”, that

¹Our study examines how to optimally combine factors in an investment portfolio. We do not address the broader question whether factor-based allocation is really superior to asset allocation. For example, it is conceivable that factors are simply a rotation of the underlying asset classes (Idzorek & Kowara, 2013).

is, the fact that when many portfolios are evaluated individually on the same data set, some are bound to show superior performance by chance alone.

By applying a multiple testing design, our study is related to Arnott, Harvey, Kalesnik, and Linnainmaa (2019), who suggest that many investors develop exaggerated expectations about factor performance as a result of data snooping. Furthermore, we contribute to the literature on factor investing by asserting that factor-based allocation is relatively easy to implement using tradable instruments; by showing that such strategies earn significant risk-adjusted excess returns relative to the market; and by demonstrating that it is difficult to beat an equally weighted factor portfolio, thereby providing guidance on the best method(s) to choose when constructing factor-based portfolios (Koedijk et al., 2016a). Finally, our analysis is closely related to the recent literature on optimization methods (Bessler & Wolff, 2015; De Miguel, Garlappi, & Uppal, 2009; Hsu, Han, Wu, & Cao, 2018; Kritzman, Page, & Turkington, 2010) and applies previously researched methods to a new data set, that is, global equity and fixed income factor premia that most investors can easily invest in. Our approach is straightforward to replicate for almost any investor and, for practical applications, can be extended in-house.

Our results confirm that factor investing, or rather factor-based allocation, is profitable, as we find positive returns in excess of cash that are robust to data-snooping corrections for some strategies. Although cash is technically speaking the correct benchmark for long-short strategies, we also apply the equity market portfolio as an alternative benchmark. Again, accounting for data-snooping biases, at least the naïve equally weighted factor portfolio is able to beat the equity market benchmark on a risk-adjusted basis. Perhaps even more important, our results ascertain that investment portfolios can be efficiently diversified using factor-based allocation because most factor portfolios deliver robust economic performance across a variety of macroeconomic states. Finally, we establish that a naïve equal weighting of factors exhibits comparatively good economic performance that cannot be outperformed by more sophisticated allocation strategies within Hansen's (2005) multiple testing framework, thus mirroring previous results for traditional asset-based allocation strategies (De Miguel et al., 2009; Hsu et al., 2018). Our main finding that the particular portfolio allocation technique does not matter much, and thus the equally weighted factor strategy is hard to beat, provides guidance for investors who are interested in implementing factor optimization (Koedijk et al., 2016a).

Our findings also reinforce the notion that factor timing is difficult, and strategic diversification across factors tends to outdo any attempts to actively time them. Both Asness (2016) and Asness, Chandra, Imanen, and Israel (2017) argue that factor timing strategies are too correlated with the basic factor strategies themselves to have a great impact on a portfolio that already includes a strategic allocation to factors. Such strategies may result in large bets on some of the factors, thereby weakening performance due to forgone diversification. In brief, lost diversification is not overcome by potential timing benefits.² Most recently, Dichtl, Drobetz, Lohre, Rother, and Vosskamp (2019) show evidence that equity factors are predictably related to fundamental and technical time-series indicators as well as factor characteristics such as factor momentum and crowding, but they conclude that such predictability is hard to take advantage of after transaction costs.

The remainder of our study is as follows. Section 2 describes the factors and factor allocation strategies we apply. Section 3 contains an empirical analysis of factor premia. Section 4 presents

²Similarly, Lee (2017) suggests that factor timers must accept some degree of lost diversification in a portfolio as potential opportunity costs of factor timing. Other authors such as Arnott, Beck, Kalesnik, and West (2016) and Bender, Sun, Thomas, and Zdorovtsov (2018) also recognize that factor prediction based on market, sentimental, and macroeconomic indicators is not an easy task. However, they do not go as far to believe it is completely futile provided that an investor has a long horizon and a good understanding of the investment rationale behind the factors.

the performance results of the ensuing factor-based portfolios and their comparisons in a multiple testing framework. Section 5 shows the results of our robustness checks. Section 6 concludes.

2 | EMPIRICAL PROCEDURE

2.1 | Factors

Ilmanen and Kizer (2012) point out that there is little consensus in the asset pricing literature over which factors should earn significant long-run rewards. Recognizing this controversial debate and the ever increasing number of candidate factors, Pukthuanthong, Roll, and Subrahmanyam (2019) introduce a thorough protocol for identifying genuine factors, that is, factors that are related to the covariance matrix of returns, are priced in their cross-section, and yield a reasonable reward-to-risk ratio.³ Using this protocol, they identify several factors, including the equity and the term premia. Moreover, they recognize several characteristic-based factors that are unrelated to risk, but correlated with average asset returns, thus contrasting market efficiency and offering potential profit opportunities. Beyond these considerations, Ang et al. (2009) emphasize that the set of factors should be parsimonious, and that factors must have the capacity to be traded in sufficiently large quantities. Taking a more practical perspective, Blyth, Szigety, and Xia (2016) stress that the chosen factors should be able to capture an investor's risk and return objectives and general investment strategy.

Taking these considerations into account, we select a large set of factors that are suitable for a US institutional investor who is allocating to global equity and fixed income securities but refrains from more sophisticated investment strategies such as merger arbitrage. Table 1 summarizes the factors and their construction based on tradable investment indices⁴ (all data underlying the factors are denominated in US dollars).

Equity. The “equity” premium is arguably the most extensively researched factor and has traditionally been the most important source of long-run excess returns (Campbell, 2008). It is well documented in international stock markets (Dimson, Marsh, & Staunton, 2018) and has been used in factor-based portfolio construction in several studies (Clarke et al., 2005; Idzorek & Kowara, 2013; Ilmanen & Kizer, 2012). In our empirical analyses, we measure the equity premium for three different regions as the return of the respective MSCI index in excess of the risk-free rate. Based on the MSCI classification, we consider three regions: the USA, EAFE (21 developed markets across Europe, Australasia, and the Far East), and EM (24 emerging markets). We measure the risk-free rate as the return on a 1-month US Treasury bill. In addition, we construct a “combined” equity premium (MKT) by equally weighting the equity premia of each region.

Value. The “value” premium was first documented by Basu (1977) in the US stock market and measures the outperformance of value stocks (i.e. stocks with comparatively low valuation

³Hsu, Kalesnik, and Viswanathan (2015) provide a more heuristic approach to factor identification that relies on the extent to which a candidate factor is debated in the academic literature, its persistence across time and geographies, and its robustness to reasonable perturbations in its definition.

⁴When constructing long-short combinations, we adjust the investment indices for time-varying beta and duration exposures (footnotes 5 and 10 below provide additional details). However, such beta and duration hedging introduces leverage in the portfolio because our strategies invest more or less in cash depending on the hedging ratio. This approach may not be appropriate for leverage-constrained investors. Therefore, in a previous version of this paper, we refrain from explicit beta or duration hedging, with qualitatively similar results.

TABLE 1 Factors

| “Single” factor | Factor construction | Data availability | “Combined” factor |
|----------------------------|---|--------------------------|------------------------------|
| MKT (EAFE) | MSCI EAFE Total Return Index (USD) minus risk-free rate | 12:1969–12:2019 | MKT |
| MKT (EM) | MSCI Emerging Markets (EM) Total Return Index (USD) minus risk-free rate | 12:1987–12:2019 | |
| MKT (US) | MSCI USA Total Return Index (USD) minus risk-free rate | 12:1969–12:2019 | |
| VAL (EAFE) | MSCI EAFE Value Total Return Index (USD) minus MSCI EAFE Growth Total Return Index (USD) | 12:1974–12:2019 | VAL |
| VAL (EM) | MSCI EM Value Total Return Index (USD) minus MSCI EM Growth Total Return Index (USD) | 12:1996–12:2019 | |
| VAL (US) | MSCI USA Value Total Return Index (USD) minus MSCI USA Growth Total Return Index (USD) | 12:1974–12:2019 | |
| CAP (EAFE) | MSCI EAFE Small Cap Total Return Index (USD) minus MSCI EAFE Large Cap Total Return Index (USD) | 12:2000–12:2019 | CAP |
| CAP (EM) | MSCI EM Small Cap Total Return Index (USD) minus MSCI EM Large Cap Total Return Index (USD) | 05:1994–12:2019 | |
| CAP (US) | MSCI USA Small Cap Total Return Index (USD) minus MSCI USA Large Cap Total Return Index (USD) | 12:2000–12:2019 | |
| MOM (EAFE) | MSCI EAFE Momentum Total Return Index (USD) minus MSCI EAFE Total Return Index (USD) | 06:1994–12:2019 | MOM |
| MOM (EM) | MSCI EM Momentum Total Return Index (USD) minus MSCI EM Total Return Index (USD) | 06:1994–12:2019 | |
| MOM (US) | MSCI USA Momentum Total Return Index (USD) minus MSCI USA Total Return Index (USD) | 06:1994–12:2019 | |
| QUA (EAFE) | MSCI EAFE Quality Total Return Index (USD) minus MSCI EAFE Total Return Index (USD) | 06:1994–12:2019 | QUA |
| QUA (EM) | MSCI EM Quality Total Return Index (USD) minus MSCI EM Total Return Index (USD) | 06:1994–12:2019 | |
| QUA (US) | MSCI USA Quality Total Return Index (USD) minus MSCI USA Total Return Index (USD) | 06:1994–12:2019 | |
| VOL (EAFE) | MSCI EAFE Minimum Volatility Index (USD) minus MSCI EAFE Total Return Index (USD) | 05:1988–12:2019 | VOL |

TABLE 1 (Continued)

| “Single” factor | Factor construction | Data availability | “Combined” factor |
|----------------------------|---|--------------------------|------------------------------|
| VOL (EM) | MSCI EM Minimum Volatility Index (USD) minus MSCI EM Total Return Index (USD) | 05:1993–12:2019 | |
| VOL (US) | MSCI USA Minimum Volatility Index (USD) minus MSCI USA Total Return Index (USD) | 05:1988–12:2019 | |
| TERM (US) | Bloomberg Barclays (BB) US Treasury Total Return Index Unhedged (USD) (LUATTRUU) minus risk-free rate | 01:1973 –12:2019 | TERM |
| TERM (Global) | BB Global Treasury Total Return Index Hedged (USD) (BTSYTRUH) minus risk-free rate | 01:1987–12:2019 | |
| REAL (US) | BB US Govt. Inflation-Linked Bonds Total Return Index (USD) (BCIT1T) minus risk- free rate | 02:1997 –12:2019 | REAL |
| REAL (Global) | BB World Govt. Inflation-Linked All Maturities Total Return Index (USD) (BCIW1T) minus risk-free rate | 12:1996–12:2019 | |
| CREDIT (US) | BB US Corporate Total Return Index Unhedged (USD) (LUACTRUU) minus LUATTRUU | 01:1973–12:2019 | CREDIT |
| CREDIT (Global) | BB Global Corporate Total Return Index Hedged (USD) (LGCPTRUH) minus BTSYTRUH | 09:2000–12:2019 | |
| HY (US) | BB US Corporate High Yield Total Return Index Unhedged (USD) (LF98TRUU) minus LUACTRUU | 07:1983–12:2019 | HY |
| HY (Global) | BB Global High Yield Total Return Index Hedged (USD) (LG30TRUH) minus LGCPTRUH | 09:2000–12:2019 | |

ratios) over growth stocks. It has subsequently been documented in international stock markets (Fama & French, 2012a) as well as across asset classes (Asness, Moskowitz, & Pedersen, 2013), and it has been included as a factor in a series of studies (Ang et al., 2009; Bender et al., 2010; Idzorek & Kowara, 2013; Ilmanen & Kizer, 2012). In our empirical tests, we construct a beta-neutral value premium for each region as long-short combinations of the respective MSCI value and growth indices.⁵ Moreover, by equally weighting the value premia of each region, we also construct a “combined” value premium (VAL).⁶

⁵Following Bender and Wang (2015), we apply a beta adjustment to the long and short positions, which we achieve by dividing the return of the short leg index by its beta and multiplying it by the beta of the long leg index. The betas of the long and short legs relative to the respective MSCI market index are computed using rolling estimation windows of 60 months.

⁶In practice, such a (theoretical) isolation of the value premium might not be attainable if liquid and cost-efficient instruments, especially on the short side, are absent. Therefore, as a robustness check, we replace the short side of the value premium by the respective MSCI market index. The ensuing results are comparable to those presented below and are available upon request.

Size. The “size” premium as a measure of excess returns of small-capitalization stocks relative to large-capitalization stocks was documented by Banz (1981) and Reinganum (1981). While there is an ongoing dispute regarding its persistence (Alquist, Israel, & Moskowitz, 2018; Horowitz, Loughran, & Savin, 2000; Schwert, 2003), many studies have taken the size premium into account when constructing factor-based portfolios (Ang et al., 2009; Bender et al., 2010; Idzorek & Kowara, 2013; Ilmanen & Kizer, 2012). We construct a beta-neutral size premium for each region as a long-short combination of the respective MSCI Small-Cap and Large-Cap indices, and also form an equally weighted “combined” size premium (CAP) using the size premia of each region (see footnote 5 for details on the beta adjustment).

Momentum. Jegadeesh and Titman (1993) show that stocks with relatively high recent returns significantly outperform stocks with relatively low recent returns. Similarly to the value premium, this “momentum” premium has also been documented in international stock markets (Fama & French, 2012a) as well as across asset classes (Asness et al., 2013), and it has been included as a factor in a number of studies (Ang et al., 2009; Bender et al., 2010; Clarke et al., 2005; Ilmanen & Kizer, 2012). We construct a beta-neutral momentum premium for each region by going long the respective MSCI Momentum index and shorting its market index in order to neutralize overall market risk (see footnote 5 for details on the beta adjustment). In addition, we construct a “combined” momentum premium (MOM) by equally weighting the momentum premia of each region.

Quality. The “quality” premium captures the outperformance of “quality” stocks, for example, firms in good financial condition in terms of profitability or leverage, over “junk” stocks (Asness, Frazzini, & Pedersen, 2019; Novy-Marx, 2014; Piotroski, 2000). While Hamdan, Pavlowsky, Roncalli, and Zheng (2016) point out that quality-based exchange-traded fund products are still scarce, several more recent studies include the quality premium as a factor (Ang, Jiang, Maloney, & Metchick, 2018; Bender et al., 2018).⁷ We construct a beta-neutral quality premium for each region by going long the respective MSCI quality index and shorting its market index (see footnote 5 for details on the beta adjustment). By equally weighting the quality premia of each region, we additionally construct a “combined” quality premium (QUA).

Low volatility. The “low-volatility” premium measures the excess returns of stocks with low volatility relative to stocks with high volatility (Ang, Hodrick, Xing, & Zhang, 2006, 2009; Haugen & Heins, 1975), and it has been included as a factor in earlier studies (Ang et al., 2018; Bender et al., 2018). We construct a beta-neutral low-volatility premium for each region by going long the respective MSCI Minimum Volatility index and shorting its market index, and also form an equally weighted “combined” low-volatility premium (VOL) using the low-volatility premia of each region (see footnote 5 for details on the beta adjustment).

Term. The “term” premium measures the performance of long-term government bonds relative to short-term government bonds, thus representing the required reward for duration extension (Ilmanen, 2011).⁸ It has been documented in international bond markets by Dimson et al. (2018), among others, and it has been used in factor-based portfolio construction

⁷Some studies, such as Asness, Ilmanen, Israel, and Moskowitz (2015) or Ang et al. (2018), use the term “defensive” premium, which can relate to either the quality or the low-volatility premium, depending on the construction.

⁸Ang et al. (2009) emphasize that the effects of macro factors, such as inflation, are reflected in fixed income prices and term spreads. Therefore, we do not consider a separate “inflation” premium as in Greenberg, Babu, and Ang (2016) or Bass, Gladstone, and Ang (2017), but instead rely on fixed income-based factors.

(Ang et al., 2009; Bender et al., 2010; Clarke et al., 2005; Idzorek & Kowara, 2013; Ilmanen & Kizer, 2012). In our empirical analyses, we follow AQR (2018) and construct both a US and a global term premium, using the respective Bloomberg Barclays (BB) Treasury bond indices in excess of the risk-free rate. In addition, we construct a “combined” term premium (TERM) by equally weighting the US and global term premia.

Real rates. The “real rates” premium captures the risk of bearing exposure to real interest rate changes and measures the performance of inflation-linked government bonds relative to nominal government bonds (Greenberg et al., 2016). While AQR (2018) suggest that investors are not equally (if at all) affected by real interest rate risk, other studies have included the real rates premium as a factor in factor-based portfolios (Bass et al., 2017; Blyth et al., 2016; Clarke et al., 2005; Greenberg et al., 2016). We construct both a US and a global real rates premium as the return of the respective BB inflation-linked bond index in excess of the risk-free rate. By equally weighting the US and global real rates premia, we additionally construct a “combined” real rates premium (REAL).

Credit. The “credit” premium measures the excess returns of investment-grade corporate bonds relative to government bonds of similar maturities, thus representing the compensation for the uncertain repayments of those bonds,⁹ and it has been included as a factor in several studies (Ang et al., 2009; Bender et al., 2010; Idzorek & Kowara, 2013; Ilmanen & Kizer, 2012; Naik, Devarajan, Nowobilski, Page, & Pedersen, 2016). We construct both a US and a global credit premium as duration-neutral long-short combinations of the respective BB corporate and Treasury bond indices,¹⁰ and form an equally weighted “combined” credit premium (CREDIT) using the US and global credit premia.

High yield. The “high-yield” premium captures the excess returns of high-yield corporate bonds relative to investment-grade corporate bonds and is primarily driven by changing default rates (Ilmanen, 2011). Several studies have taken the high-yield premium into account when constructing factor-based portfolios (Ang et al., 2009; Bender et al., 2010; Blyth et al., 2016; Naik et al., 2016). In the empirical analyses, we construct both a US and a global duration-neutral high-yield premium by going long the BB high-yield and shorting the associated corporate bond indices.¹¹ We also construct a “combined” high-yield premium (HY) by equally weighting the US and global high-yield premia.

2.2 | Factor allocation strategies

In order to identify the “best” factor allocation strategy, we are guided by the literature on asset allocation (e.g. De Miguel et al., 2009; Hsu et al., 2018). In total, we construct 17 factor allocation strategies, listed in Table 2, that either are estimation-free or require the estimation of some factor moments.

⁹Yield spreads over government bonds likely also reflect lower liquidity of corporate bonds (Ilmanen, 2011).

¹⁰Following Maeso, Martellini, and Rebonato (2019), we apply a duration adjustment to the long and short positions, which we achieve by dividing the return of the short-leg bond index by its duration and multiplying it by the duration of the long-leg bond index. Duration adjustment achieves approximately equal risk between the long and short positions. We use modified duration data from Bloomberg. The average duration over the sample period is 6.6 (5.9) years for the US (global) credit premium.

¹¹See footnote 10 for details on the duration adjustment. The average duration over the sample period is 4.3 (4.6) years for the US (global) high-yield premium.

TABLE 2 Factor allocation strategies

This table reports the factor allocation strategies considered together with their abbreviations used throughout this study and indicates whether they require an estimation of the expected covariance matrix $\hat{\Sigma}_t$ and possibly the expected factor returns $\hat{\mu}_t$ (refer to Section 2.2 for details about the construction).

| Strategy | $\hat{\Sigma}_t$ | $\hat{\mu}_t$ | Abbreviation |
|--|------------------|---------------|-----------------------------------|
| (a) Naïve allocation | | | |
| Equally weighted portfolio | — | — | EW |
| (b) Risk-based strategies | | | |
| Minimum-variance portfolio | + | — | MinVar |
| Equally weighted risk contribution | + | — | ERC |
| Most diversified portfolio | + | — | MD |
| Volatility timing portfolio | + | — | VT |
| (c) Risk- and return-based strategies | | | |
| Reward-to-risk timing portfolio | + | + | RRT (mean), RRT (EP), RRT (TSMOM) |
| Mean-variance optimization | + | + | MV (mean), MV (EP), MV (TSMOM) |
| Bayes–Stein shrinkage approach | + | + | BS (mean), BS (EP), BS (TSMOM) |
| Black–Litterman model | + | + | BL (mean), BL (EP), BL (TSMOM) |
| (d) Mixed strategies | | | |
| Equally weighted combination of EW and MinVar | + | — | Mixed (EW & MinVar) |
| Equally weighted combination of EW and MV (mean) | + | + | Mixed (EW & MV) |
| Equally weighted combination of MinVar and MV (mean) | + | + | Mixed (MinVar & MV) |

2.2.1 | Naïve allocation

Equally weighted portfolio. A popular allocation strategy is the naïve $1/N$ rule that assigns equal, time-invariant weights, that is, $w_{it} = 1/N$, to each factor. This rule, denoted as EW, is the easiest to implement as it does not require any estimation of factor moments.

2.2.2 | Risk-based allocation strategies

Strategies that base their allocation decisions only on risk measures without considering expected returns have become increasingly popular among investors. These risk-based allocation strategies rely on the premise that expected (co)variances can be estimated with greater precision than expected returns (Chopra & Ziemba, 1993; Merton, 1980), and thus one way to reduce estimation risk is to refrain from estimating expected returns altogether. All risk-based allocation strategies have in common that their weights are a decreasing function of risk (Clarke, de Silva, & Thorley, 2013).

Minimum-variance portfolio. The minimum-variance (MinVar) portfolio minimizes the variance of portfolio returns:

$$\min_{w_t} w_t' \hat{\Sigma}_t w_t, \quad (1)$$

where w_t denotes the vector of portfolio weights invested in N factors, and $\hat{\Sigma}_t$ is an estimate of the expected covariance matrix. The solution to this performance by chance alone is given by

$$w_t = \frac{\hat{\Sigma}_t^{-1} \mathbf{1}_N}{\mathbf{1}_N' \hat{\Sigma}_t^{-1} \mathbf{1}_N},$$

where $\mathbf{1}_N$ denotes a vector of N ones.

For all allocation strategies, we estimate $\hat{\Sigma}_t$ as sample covariance matrices using a rolling estimation window of 60 months, include a budget restriction to ensure that the portfolio weights sum to 1, that is, $w_t' \mathbf{1}_N = 1$, and exclude short-selling, that is, $w_{it} \geq 0$. As emphasized by Idzorek and Kowara (2013), although we constrain the weights to be non-negative, given the construction of the factors, factor-based allocation strategies do require short-selling.

Equally weighted risk contribution portfolio. Building on earlier work by Qian (2006, 2011), Maillard, Roncalli, and Teiletche (2010) develop a heuristic approach that aims at equalizing the risk contribution (RC) from each factor in a portfolio. The RC of a factor, $\sigma_i(w)$, is the share of total portfolio risk attributable to that factor, and is computed as the product of the respective portfolio weight w_i and its marginal risk contribution (MRC), $\delta_{w_i} \sigma(w)$, where the MRC measures the change in total portfolio risk induced by an incremental increase in the respective portfolio weight, that is,

$$\delta_{w_i} \sigma(w) = \frac{\delta \sigma(w)}{\delta w_i} = \frac{w_i \sigma_i^2 + \sum_{j \neq i} w_j \sigma_{ij}}{\sigma(w)}.$$

Because the equally weighted risk contribution (ERC) portfolio compares the relative risk of a factor to the rest of the portfolio, the solution is not in closed form, but must be found numerically. Maillard et al. (2010) recommend using a sequential quadratic programming (SQP) algorithm to solve an optimization problem, which effectively minimizes the variance of the (rescaled) risk contributions:

$$\min_{w_t} \sum_{i=1}^N \sum_{j=1}^N \left(w_{it} (\hat{\Sigma}_t w_t)_i - w_{jt} (\hat{\Sigma}_t w_t)_j \right)^2. \quad (2)$$

Maillard et al. (2010) further show that the ERC portfolio coincides with the equally weighted portfolio if all factors exhibit the same volatility and uniform cross-correlations. Under the assumption of uniform cross-correlations (but different volatilities), the ERC portfolio weights are proportional to the inverse of the factors' individual (estimated) volatilities, that is,

$$w_{it} = \frac{1/\hat{\sigma}_{it}}{\sum_{i=1}^N (1/\hat{\sigma}_{it})}.$$

Most diversified portfolio. Choueifaty and Coignard (2008) develop an allocation strategy that uses diversification as the criterion in portfolio construction by maximizing the diversification ratio (DR), that is, the ratio of the weighted average of volatilities to total portfolio volatility,

$$\max_{w_t} \frac{w_t' \hat{\sigma}_t}{\sqrt{w_t' \hat{\Sigma}_t w_t}}, \quad (3)$$

where $\hat{\sigma}_t$ is a vector of estimated standard deviations at time t . The most diversified (MD) portfolio simply maximizes the ratio between two distinct definitions of portfolio volatility, that is, the ratio between the average portfolio factors' volatility and the total portfolio volatility. The solution to this optimization problem is found numerically by using a SQP algorithm.

The DR measures the degree of diversification of a portfolio. The higher the DR, the lower is the risk of the factor-based portfolio $\sqrt{w_t' \hat{\Sigma}_t w_t}$ compared to the combination of risks of the individual factors $w_t' \hat{\sigma}_t$, provided that the factors are not perfectly correlated. Choueifaty and Coignard (2008) show that the MD portfolio coincides with the minimum-variance portfolio if all factors exhibit the same volatility, and with the equally weighted risk contribution portfolio under the assumption of uniform cross-correlations. Moreover, Choueifaty, Froidure, and Reynier (2013) demonstrate that under the assumption of a homogenous investment universe, that is, assuming that each factor's expected excess return is proportional to its volatility, and thus inferring identical ex-ante Sharpe ratios across factors, maximizing the DR is equivalent to maximizing the Sharpe ratio.¹²

Volatility timing portfolio. Fleming, Kirby, and Ostdiek (2001, 2003) investigate a volatility timing (VT) strategy that rebalances the portfolio weights within a mean–variance optimization framework based only on expected volatility changes, while treating expected returns as constant. Kirby and Ostdiek (2012) extend this strategy by, *inter alia*, allowing the sensitivity of the weights against those volatility changes to be adjusted in order to control turnover and transaction costs. Moreover, they assume that all pairwise correlations between the factor returns are zero, that is, $\hat{\Sigma}_t$ is a diagonal matrix, in order to further reduce estimation risk. Therefore, the weights of the VT portfolio are given by

$$w_{it} = \frac{(1/\hat{\sigma}_{it}^2)^\eta}{\sum_{i=1}^N (1/\hat{\sigma}_{it}^2)^\eta}, \quad (4)$$

where $\hat{\sigma}_{it}^2$ is the estimated variance of factor i at time t . The tuning parameter η measures how aggressively an investor adjusts portfolio weights in response to volatility changes. Kirby and Ostdiek (2012) point out that, while setting $\eta = 0$ corresponds to the equally weighted portfolio, increasing $\eta \rightarrow \infty$ increases the weight on the factor with the lowest volatility toward 1. In our empirical analyses, we set $\eta = 1$, which is equivalent to the minimum-variance portfolio under the assumption that factors are uncorrelated (Roncalli, 2013).

¹²Given that we also consider genuine risk premia such as the equity premium in our factor set, the assumption of homogenous ex-ante Sharpe ratios across factors is not met for our analysis.

2.2.3 | Risk- and return-based allocation strategies

All risk-based allocation strategies discussed above ignore information on expected factor returns. As a matter of course, we next analyze whether their performance can be improved by incorporating such information, while accounting for the possibility of estimation errors in returns.

Reward-to-risk timing portfolio. The reward-to-risk timing (RRT) portfolio aims to improve upon the VT portfolio (Kirby and Ostdiek, 2012). In particular, the RRT strategy assigns greater weight to factors with comparatively high expected return and low expected variance, and vice versa. Formally, assuming again that all pairwise correlations are zero, the weights of the RRT portfolio are given by

$$w_{it} = \frac{(\hat{\mu}_{it}^+ / \hat{\sigma}_{it}^2)^\eta}{\sum_{i=1}^N (\hat{\mu}_{it}^+ / \hat{\sigma}_{it}^2)^\eta}, \quad (5)$$

with $\hat{\mu}_{it}^+ = \max(\hat{\mu}_{it}, 0)$, where $\hat{\mu}_{it}$ denotes the estimated return of factor i at time t , that is, the investor has strong prior belief that $\mu_{it} \geq 0$ for all factors and eliminates any factor with $\hat{\mu}_{it} < 0$ from consideration at time t . Again, we set the tuning parameter $\eta = 1$.

We use three different estimates for the vector of expected factor returns μ_t . In line with Idzorek and Kowara (2013), our first estimate of $\hat{\mu}_t$ is based on historical returns. In particular, we take sample means based on a 60 month rolling window and denote the ensuing portfolio as RRT (mean).

Building on findings by Bender et al. (2018) that point toward some predictive information in macroeconomic and financial variables for future factor returns,¹³ we provide a second estimate of $\hat{\mu}_t$ using economic predictors and denote the resulting portfolio by RRT (EP). Our approach is as follows. In a first step, we estimate a univariate predictive regression using rolling estimation windows of 60 months for each factor i :

$$r_{i,t+1} = \alpha + \beta x_t + \varepsilon_{i,t+1}, \quad (6)$$

where $r_{i,t+1}$ is the return of factor i from period t to $t + 1$, x_t a predictor variable known at time t , and $\varepsilon_{i,t+1}$ a zero-mean disturbance term. As independent variables, we follow Arnott, Kelso, Kiscadden, and Macedo (1989) and Bender et al. (2018) and use a small set of macroeconomic and financial indicators. We consider the year-on-year (YOY) inflation rate (as measured by the Consumer Price Index (CPI) and the Producer Price Index (PPI)), the YOY change in M1 and M2 money supply, the YOY change in personal income and consumption expenditures, and the annual personal saving rate.¹⁴

To justify the choice of these macroeconomic variables, we test their in-sample predictive power. When regressing factor returns on lagged predictors (not reported), the R^2 values are low, ranging from 0% to slightly more than 10%. These low values are consistent

¹³Additional empirical evidence is provided by Chincoli and Guidolin (2017) on the predictability of equity style returns (e.g. size or value premia), and Ludvigson and Ng (2009) on the predictability of bond risk premia.

¹⁴These macroeconomic variables are also used as part of a large set of predictors in Ludvigson and Ng (2009) to forecast both equity and bond premia. All data series are retrieved from the Federal Reserve Bank of St. Louis (<https://fred.stlouisfed.org>).

with prior studies; see, for example Dichtl, Drobetz, Neuhierl, and Wendt (2020) for a review of the stock prediction literature. However, Campbell and Thompson (2008) show that even small R^2 statistics can generate large benefits for investors, thus one should expect predictive regressions to have only modest explanatory power. Conversely, they claim that regressions with large statistics would be too profitable to believe. Furthermore, as documented by Leitch and Tanner (1991), there is only a weak association between statistical measures of forecasting performance and economic forecast profitability. Strategies that outperform the benchmark model in terms of statistical measures often fail to outperform when using profit- or utility-based metrics.

In a second step, we compute the predicted factor returns by using the fitted model from Eq. (6) in conjunction with the realized value of x_t to construct the 1-month-ahead forecast $\hat{r}_{i,t+1}$. We then average the individual forecasts based on the K different independent variables, that is, we compute a combination forecast

$$\hat{r}_{i,t+1,\text{mean}} = \sum_{k=1}^K \frac{\hat{r}_{i,k,t+1}}{K}. \quad (7)$$

in order to capture varying degrees of adaptability to structural breaks and mitigate the problem of potential model misspecification (Timmermann, 2006).

Finally, we provide a third estimate of $\hat{\mu}_t$ by exploiting each factor's time-series momentum (TSMOM), drawing on the recent findings by Gupta and Kelly (2019), who show that individual factors can be timed based on their own recent performance.¹⁵ The resulting portfolio is denoted by RRT (TSMOM). Specifically, we first construct a TSMOM indicator for each factor i that equals 1 if the factor exhibits positive TSMOM, that is, the cumulative return over the past $m = \{1, 3, 6, 9, 12\}$ months is equal to or greater than zero, and 0 otherwise. We then follow Neely, Rapach, Tu, and Zhou (2014) and transform the TSMOM indicators into point forecasts of the factor returns by using the respective indicator as an independent variable in the predictive regression model in Eq. (6). Predicted factor returns are again computed by constructing the 1-month-ahead forecasts $\hat{r}_{i,t+1}$, and averaging across the K different forecasts (see Eq. [7]).

Mean-variance optimization. Following the mean-variance (MV) approach developed by Markowitz (1952), we solve the following optimization problem:

$$\max_{w_t} w_t' \hat{\mu}_t - \frac{\gamma}{2} w_t' \hat{\Sigma}_t w_t, \quad (8)$$

where w_t denotes the vector of portfolio weights invested in N factors, $\hat{\mu}_t$ is an estimate of the vector of expected factor returns, $\hat{\Sigma}_t$ is an estimate of the expected covariance matrix, and γ measures the coefficient of relative risk aversion. We assume a risk aversion coefficient of $\gamma = 2$ (De Miguel et al., 2009). As in the RRT strategy, we use three estimates of $\hat{\mu}_t$ based on sample means (MV (mean)), economic predictors (MV (EP)), and time-series momentum (MV (TSMOM)).

Bayes-Stein shrinkage approach. The Bayes-Stein (BS) shrinkage approach was proposed by Jorion (1985, 1986), and is one of the most prominent extensions of the MV approach. In an attempt to reduce estimation errors, the BS approach employs the idea of shrinkage estimation due to Stein (1956) and James and Stein (1961). In line with Hsu et al. (2018), we extend this

¹⁵Huang, Jiang, Tu, and Zhou (2017) also show that the returns of portfolios sorted on cross-sectional characteristics such as size and value can be predicted based on the past 12-month excess stock market return (TSMOM).

methodology and shrink $\hat{\mu}_t$ toward the implied return of the equally weighted portfolio, that is, $\hat{\mu}_t^{\text{ew}} = \gamma \hat{\Sigma}_t w_t^{\text{ew}}$, as follows:

$$\hat{\mu}_t^{\text{BS}} = (1 - \hat{\lambda}_t) \hat{\mu}_t + \hat{\lambda}_t \hat{\mu}_t^{\text{ew}}, \text{ with } \hat{\lambda}_t = \frac{N + 2}{(N + 2) + T(\hat{\mu}_t - \hat{\mu}_t^{\text{ew}})' \hat{\Sigma}_t^{-1} (\hat{\mu}_t - \hat{\mu}_t^{\text{ew}})}, \quad (9)$$

where T is the estimation window length (i.e. 60-month rolling windows in our analyses). Again, we assume a risk aversion coefficient of $\gamma = 2$.¹⁶ Analogously to the RRT strategy, we provide three estimates of $\hat{\mu}_t$ based on sample means (BS (mean)), economic predictors (BS (EP)), and time-series momentum (BS (TSMOM)). In addition, an estimate of the covariance matrix is provided as:

$$\hat{\Sigma}_t^{\text{BS}} = \frac{1}{T - N - 2} \sum_{s=t-T}^{t-1} (R_s - \hat{\mu}_t)(R_s - \hat{\mu}_t)'. \quad (10)$$

The adjusted estimates $\hat{\mu}_t^{\text{BS}}$ and $\hat{\Sigma}_t^{\text{BS}}$ are then used in the MV optimization problem (Eq. (8)) to derive the portfolio weights.

Black–Litterman model. The Black–Litterman (BL) model, advocated by Black and Litterman (1992), enables investors to combine their “subjective” views on expected factor returns with those implied by a “reference” portfolio that represents the desired allocation of factors in the absence of views. By incorporating the confidence in their views, investors can control how strongly the views influence the portfolio weights. Therefore, investors will only depart from the reference portfolio if they are able to provide reliable estimates of expected factor returns.

The implied factor returns Π_t are derived using reverse optimization, that is, assuming that the reference weights w_t^* are the result of a MV optimization:

$$\Pi_t = \gamma \hat{\Sigma}_t w_t^*. \quad (11)$$

In our empirical analyses, we use the EW portfolio as the reference portfolio, that is, $w_{it}^* = 1/N$. Subsequently, the BL master formula combines Π_t with the investor's views stacked in the vector Q_t and incorporates the uncertainty of each view as quantified in a matrix Ω_t :

$$\hat{\mu}_t^{\text{BL}} = \left[(\tau \hat{\Sigma}_t)^{-1} + P_t' \Omega_t^{-1} P_t \right]^{-1} \left[(\tau \hat{\Sigma}_t)^{-1} \Pi_t + P_t' \Omega_t^{-1} Q_t \right]. \quad (12)$$

In our analyses, P_t is a diagonal matrix containing the information for which factor a view exists, and Q_t are the views. As views, we use the previously discussed estimates of $\hat{\mu}_t$ based on sample means (BL (mean)), economic predictors (BL (EP)), and time-series momentum (BL (TSMOM)). Moreover, we follow Meucci (2010) and set

$$\Omega_t = P_t' \left(\frac{1}{c} \hat{\Sigma}_t \right) P_t,$$

¹⁶The shrinkage coefficient $\hat{\lambda}_t$ varies between 0.1 and 0.5 in our analysis.

assuming that the reliability of the views is proportional to the reliability of implied returns with a constant proportionality factor $1/c$, which we set to 1 in our analysis.¹⁷ The parameter τ controls how much the optimized portfolio may depart from the reference portfolio and is set to $\tau = 1/T$, that is, $\tau \approx 0.0167$ in our application.

Moreover, we derive the posterior covariance matrix following the approach in Satchell and Scowcroft (2000):

$$\hat{\Sigma}_t^{\text{BL}} = \Sigma_t + \left[(\tau \Sigma_t)^{-1} + P_t' \Omega_t^{-1} P_t \right]^{-1}. \quad (13)$$

We then use the combined return estimates $\hat{\mu}_t^{\text{BL}}$ together with the posterior covariance matrix $\hat{\Sigma}_t^{\text{BL}}$ in the MV optimization problem (Eq. (8)) to derive the portfolio weights.

2.2.4 | Mixed strategies

Finally, we consider “mixed” strategies that are themselves combinations of other strategies. The intuition behind this approach is that, in the presence of parameter uncertainty, all optimized portfolios are obtained with estimation errors. Given that those estimation errors are not perfectly correlated, combining any two strategies can help to diversify estimation risk (Kan & Zhou, 2007; Tu & Zhou, 2011). Moreover, as pointed out by De Miguel et al. (2009), such mixed strategies apply the idea of shrinkage directly to the portfolio weights instead of the factor moments, as in the case of BS, for example, which facilitates the selection of a specific target toward which one shrinks a given portfolio.

In our empirical analyses, we follow the “diversifying the diversifier” approach of Amenc, Goltz, Lodh, Martellini, and Shirbini (2012) and construct equally weighted combinations across two strategies each. In particular, we consider combinations of the EW and the MinVar portfolio (De Miguel et al., 2009), the EW and MV (mean) portfolio, and the MinVar and MV (mean) portfolio (Kan & Zhou, 2007).

3 | EMPIRICAL ANALYSIS OF FACTORS

The data used to construct the single factors are obtained from MSCI and Bloomberg. Due to data availability, the sample period in our analyses is from January 2001 to December 2019. After considering an initial period of 60 months for estimating sample means and covariance matrices, we analyze the performance of the factor allocation strategies from January 2006 to December 2019.

¹⁷ $c \rightarrow \infty$ implies absolute certainty, while $c \rightarrow 0$ increases the level of uncertainty that the investor assigns to her views. In a previous version of the paper, we set $c = 1/\tau$, such that the uncertainty structure is proportional to the uncertainty of implied factor returns (Idzorek, 2007). This approach resulted in comparatively small uncertainty terms, thus biasing the combined return estimates toward the investor's views and leading to portfolio weights that were comparable to those obtained from a mean–variance optimization. Following the suggestion of an anonymous referee, we increased the level of uncertainty to bias the combined return estimates more toward the implied factor returns.

3.1 | Descriptive statistics

Table 3 summarizes (arithmetic) average annualized factor returns and volatilities, as well as raw Sharpe ratios (Ang, 2014), measured as the ratio of factor return to volatility, and maximum drawdown, defined as the largest cumulative loss from peak to trough, over the sample period. In line with the factors' intuition of long-run return drivers, we document positive average factor returns (with the exception of the EAFE and US value premium).¹⁸ Unsurprisingly, the MKT premium, which is not a long-short portfolio but rather a genuine risk premium, shows the highest factor return, and it is also the riskiest factor in terms of both volatility and drawdown. On a risk-adjusted basis, the VOL premium exhibits the best performance with a Sharpe ratio of 0.95, whereas the VAL premium falls behind the other factors with a Sharpe ratio of -0.01 , delivering the lowest factor return accompanied by comparatively high risk (Panel B).

While the performance of all combined factors is positive on average, Figure 1 highlights that there is substantial variation in the performance over time and across factors. In line with the descriptive statistics in Table 3, the MKT premium exhibits the highest variation, ranging from an annual return of -45.4% in 2008 to 44.7% in 2009. Figure 1 also provides an indication regarding the correlation structure of the factors. For example, the MKT and VOL premia tend to move in opposite directions. Such diversification benefits are indispensable for constructing factor-based portfolios.

Figure 2 provides the sample correlation structure. The MKT and HY premia exhibit the highest correlations with other factors. The other factors are less correlated (if at all), and the correlation coefficients and directions are generally in line with previous studies, for example, the negative relation between the VAL and MOM premia (Asness, 1997; Bender et al., 2010). Figure 2 further highlights the strong covariation of each factor across different regions, which is in line with previous studies (Asness et al., 2013), while different factors within one region are less strongly correlated.

So far, our findings support the idea that diversification across different factors leads to improved portfolio performance due to positive average factor returns and relatively low correlations across factors. However, when assessing the potential of factor-based allocation strategies, an obtrusive question is whether the factors, given that they are constructed from tradable investment indices, are indeed uncorrelated with asset class returns. Moreover, while some factors are thought to be driven by behavioral biases and are potentially unrelated to the broader economic development, other factors seem to offer compensation for potential losses during "bad" times (Ang, 2014). Therefore, with the objective of constructing factor-based portfolios that perform well under as many scenarios as possible, we next shed light on the factors' performance conditional on asset class returns and different macroeconomic environments.

3.2 | Conditional analyses

The equity and, to a lesser extent, the term premium can be interpreted as "traditional" asset class proxies, that is, for global equities and bonds. The question whether the other factors exhibit a linear relationship with those asset classes or rather are "dynamic" in the sense of being uncorrelated with asset class returns (Fung & Hsieh, 1997) is crucial when assessing the

¹⁸The observed negative value premium is specific to the chosen sample period. Similar results are presented by, for example, Wellington (2016) over the 10-year period ending in 2015. Fama and French (2012b) estimate that the probability of a negative average value premium for a 10-year period is about 11.5%.

TABLE 3 Descriptive statistics

This table reports descriptive statistics for cash and all factors over the sample period from January 2001 to December 2019: annualized arithmetic mean of monthly returns (Mean), annualized standard deviation (SD), Sharpe ratio (SR), and maximum drawdown (MDD). MDD is defined as the largest cumulative loss from peak to trough, where the trough is the lowest value between two consecutive peaks. The factor abbreviations are explained in Table 1.

| Factor | Mean | SD | SR | MDD |
|---------------------------|-------|-------|-------|--------|
| Cash | 1.4% | 0.4% | – | 0.0% |
| Panel A: Single factors | | | | |
| MKT (EAFE) | 4.6% | 16.4% | 0.28 | –57.3% |
| MKT (EM) | 10.0% | 21.4% | 0.47 | –62.2% |
| MKT (US) | 6.4% | 14.5% | 0.44 | –51.6% |
| VAL (EAFE) | –0.3% | 5.9% | –0.05 | –17.6% |
| VAL (EM) | 0.9% | 5.7% | 0.16 | –14.4% |
| VAL (US) | –0.9% | 7.2% | –0.12 | –24.1% |
| CAP (EAFE) | 4.5% | 6.6% | 0.67 | –22.0% |
| CAP (EM) | 1.2% | 7.3% | 0.16 | –27.5% |
| CAP (US) | 2.6% | 8.9% | 0.29 | –20.2% |
| MOM (EAFE) | 1.9% | 7.1% | 0.27 | –24.8% |
| MOM (EM) | 2.6% | 6.8% | 0.38 | –26.4% |
| MOM (US) | 2.7% | 7.3% | 0.38 | –24.2% |
| QUA (EAFE) | 2.4% | 4.4% | 0.56 | –5.1% |
| QUA (EM) | 1.7% | 4.0% | 0.42 | –7.5% |
| QUA (US) | 1.8% | 3.2% | 0.57 | –10.9% |
| VOL (EAFE) | 4.0% | 4.8% | 0.84 | –8.1% |
| VOL (EM) | 3.3% | 3.7% | 0.89 | –4.8% |
| VOL (US) | 2.7% | 4.4% | 0.62 | –8.4% |
| TERM (US) | 2.8% | 4.4% | 0.64 | –5.6% |
| TERM (Global) | 3.0% | 2.9% | 1.03 | –4.5% |
| REAL (US) | 3.8% | 5.9% | 0.65 | –12.7% |
| REAL (Global) | 3.9% | 5.2% | 0.75 | –10.2% |
| CREDIT (US) | 1.2% | 4.7% | 0.24 | –26.2% |
| CREDIT (Global) | 1.5% | 3.3% | 0.45 | –15.8% |
| HY (US) | 4.2% | 7.8% | 0.54 | –28.7% |
| HY (Global) | 4.3% | 7.4% | 0.59 | –26.9% |
| Panel B: Combined factors | | | | |
| MKT | 7.0% | 16.4% | 0.43 | –57.0% |
| VAL | –0.1% | 5.0% | –0.01 | –10.9% |
| CAP | 2.7% | 5.4% | 0.50 | –16.1% |
| MOM | 2.4% | 5.9% | 0.41 | –22.6% |
| QUA | 2.0% | 2.8% | 0.71 | –4.9% |
| VOL | 3.3% | 3.5% | 0.95 | –4.9% |
| TERM | 2.9% | 3.5% | 0.81 | –4.8% |
| REAL | 3.9% | 5.4% | 0.72 | –11.1% |
| CREDIT | 1.3% | 3.9% | 0.34 | –21.1% |
| HY | 4.3% | 7.5% | 0.57 | –27.8% |

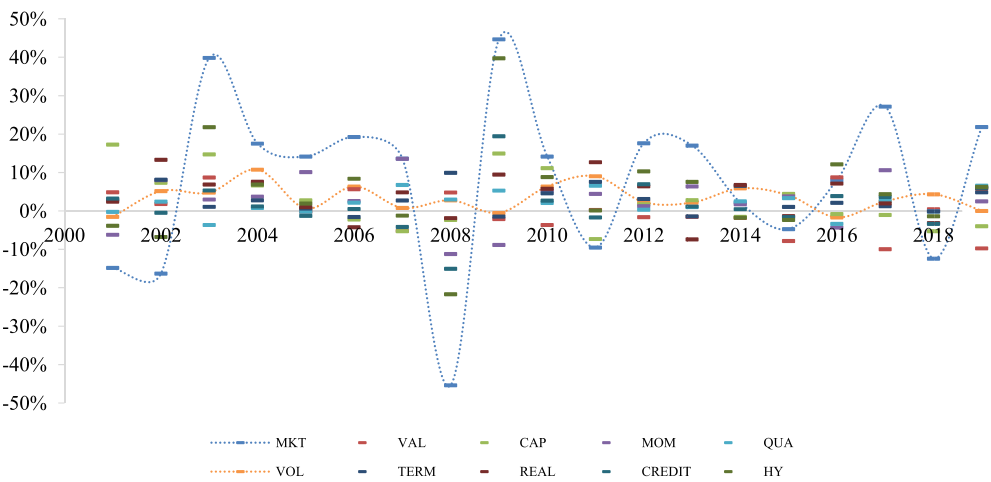


FIGURE 1 This figure depicts the annual “combined” factor returns over the sample period from January 2001 to December 2019. The factor abbreviations are explained in Table 1

potential of factor-based allocation strategies. To address this question, we follow Fung and Hsieh (1997) and sort the monthly equity and term premia into five quintiles – from severe downturns (Q1) to extreme upturns (Q5) – and compute the average factor returns in each quintile. Figure 3a,b depicts the results.

In addition, to assess whether the performance of the factors is persistent across different macroeconomic environments or driven by the compensation for “bad” times, we follow Ilmanen, Maloney, and Ross (2014) and divide our sample period into four different

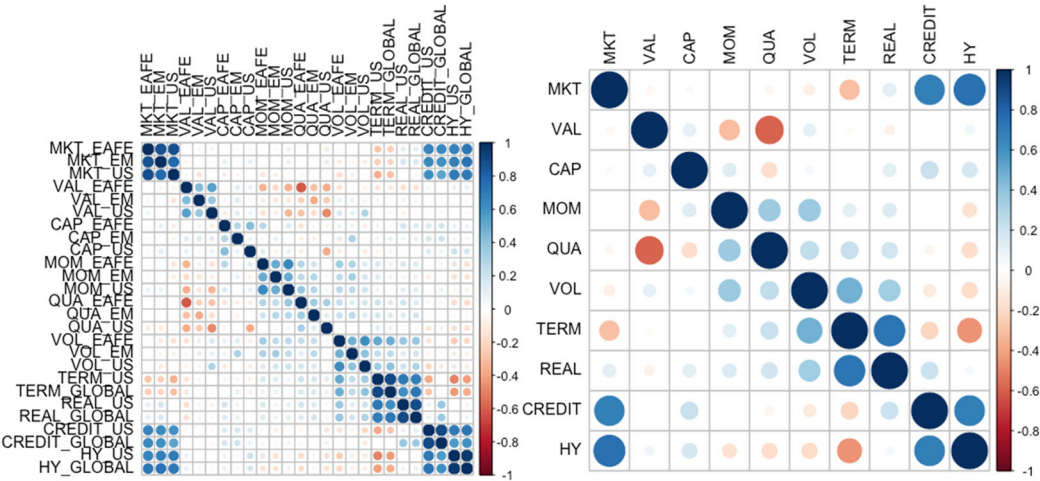


FIGURE 2 This figure depicts the matrix of correlations between the monthly factor returns over the sample period from January 2001 to December 2019 (left: single factors; right: combined factors). The factor abbreviations are explained in Table 1



FIGURE 3 This figure depicts the average monthly “combined” factor returns over the sample period from January 2001 to December 2019 over quintiles of the equity premium (MKT) in panel A, over quintiles of the term premium (TERM) in panel B, and over four different macroeconomic states in panel C. In panels A and B, Q1 (Q5) denotes the quintile with the lowest (highest) premium. In panel C, a period of low (high) growth is characterized by the Chicago Fed National Activity Index (CFNAI) below (above) its sample median, and a period of low (high) inflation is characterized by the year-on-year inflation rate (CPIYOY) below (above) its sample median. The factor abbreviations are explained in Table 1

macroeconomic “states”. We classify the states using the interaction of a growth and an inflation indicator. In particular, we use the Chicago Fed National Activity Index (CFNAI) to proxy for economic growth, and the year-on-year inflation rate (CPIYOY) to proxy for inflation.¹⁹ The “low growth, high inflation” state, for example, is characterized by below-median economic growth and above-median inflation, entailing a particularly unfavorable investment environment. Figure 3c shows the average factor returns in each state.

The results in Figure 3a show that the TERM, CREDIT, and HY premia as well as the VOL premium are highly sensitive, in opposite directions, to the movements in global equities. The remaining factors exhibit no clear linear relationship, with the results for the VAL premium partly in contrast to prior findings, for example, Fung and Hsieh (1997), who demonstrate that a value strategy is comparable to a buy-and-hold strategy in equities, which would imply a strictly linear relationship between VAL and MKT. While we observe the expected linear relationship from the second to the fifth quintile, there is a small positive (as opposed to more negative) value premium in severe downturns (Q1). Our (unreported) results indicate that this effect is driven by the short side of the portfolio, which loses even more than the long side in this state. These results might be reconciled with Zhang (2005), who shows that the value spread is countercyclical, that is, value is riskier than growth during bad times. With respect to global bonds, the results shown in Figure 3b are ambiguous. Linear relationships are discernible for the QUA, VOL, REAL, and HY premia.

The results in Figure 3c indicate that, in line with findings for global equities (Ilmanen et al., 2014), the MKT premium is especially hurt in a stagflationary (low growth, high inflation) environment. Similar results are obtained for the HY and, to a lesser degree, the CREDIT premium, which is expected given that corporate bonds and equities are contingent claims to the same firm value (Chen, Collin-Dufresne, & Goldstein, 2009). In contrast, defensive factors such as QUA excel well in this “complicated” economic environment, while falling behind in more favorable states. For the other factors, we do not find any evident pattern across macroeconomic environments, thus their performance seems to be persistent across time.

Overall, our results indicate that the factors exhibit varying asset class as well as macroeconomic sensitivities. Therefore, diversifying across factors within a portfolio might enable an investor to profit in all “states” of the world, that is, possibly generate absolute returns. However, provided that some factors (i.e. the MKT and HY premia) exhibit high correlations with other factors (see Figure 2), using naïve diversification strategies such as simple equally weighted factor portfolios might not necessarily be the most sensible approach. Instead, a conservative use of the information contained in the correlation structure may provide benefits if reasonably applied such that overfitting and excessive turnover are prevented (Ilmanen & Kizer, 2012).

4 | EMPIRICAL ANALYSIS OF FACTOR ALLOCATION STRATEGIES

To get a first grasp of the diversification and turnover properties of the different factor allocation strategies, we plot the development of their resulting “combined” factor weights over the evaluation period from January 2006 to December 2019 in Figure 4. Unsurprisingly, those strategies that rely on some form of expected return estimates, especially MV optimization and BS shrinkage, exhibit

¹⁹The CFNAI is obtained from the Federal Reserve Bank of Chicago (<https://www.chicagofed.org/research/data/cfnai/current-data>), while the CPIYOY is provided by the US Bureau of Labor Statistics and retrieved from the Federal Reserve Bank of St. Louis (<https://fred.stlouisfed.org/series/CPIAUCSL>).

more extreme portfolio reallocations, periodically even to the extent of being invested in only a single factor, which should be seen as an extreme form of factor timing. All other strategies display more balanced allocations, potentially resulting in lower portfolio turnover. Moreover, there seems to be a structural break in the factor risk structure around the global financial crisis, when all strategies display a striking shift in their portfolio allocations, primarily reallocating more toward “safer” factors, for example, the QUA, VOL, and REAL premia.

4.1 | Performance statistics

Building on the factor allocations shown in Figure 4, we construct total return portfolios by adding the risk-free rate to the (optimized) factor portfolio return, that is, we hold a 100% position in cash together with a zero-dollar position in the factor portfolio (Idzorek & Kowara, 2013; Ilmanen & Kizer, 2012). Therefore, technical speaking, cash is the proper benchmark for our long-short strategies. Because our set of factors contains genuine risk premia over cash (i.e. equity and bond premia), we additionally use the MKT portfolio as an alternative benchmark.²⁰ Round-trip transaction costs of 50 bps are considered for all total return portfolios.²¹

Table 4 summarizes the performance statistics of a risk-free investment, the MKT portfolio, and all total return portfolios (net of transaction costs) over the evaluation period from January 2006 to December 2019. The simple EW portfolio delivers comparatively good risk-adjusted performance, regardless of whether “single” factors, that is, individual factors in each region (annualized Sharpe ratio of 0.76 in Panel A), or “combined” factors, that is, factors pooled across regions (annualized Sharpe ratio of 0.79 in Panel B), are used to construct the factor portfolio. These Sharpe ratios are even substantially higher than that of the equity market portfolio (0.12). As expected, the risk-minimization factor strategies, that is, MinVar, ERC, MD, and VT, exhibit the lowest risk in terms of volatility and maximum drawdown, and are also the most diversified ones, as indicated by the high average diversification ratio (i.e. the ratio of the weighted average of volatilities to total portfolio volatility; see Eq. (3)). At the same time, these strategies yield comparatively low average returns.²²

Turning to those strategies that also require estimates of expected factor returns, we note that the MV and BS strategies boast the highest average returns when allocating to “single” factors (Panel A of Table 4), but yield the lowest average returns when allocating to “combined” factors (Panel B of Table 4). Regardless of the factor set, however, these strategies exhibit the highest risk in terms of volatility and maximum drawdown, as well as the highest portfolio turnover. Their average diversification ratios are low, implying that these strategies are not well diversified. Moreover, the performance of these strategies is highly dependent on the type of factor return estimates, with those strategies that utilize economic predictions, for example, MV (EP), doing comparatively well.

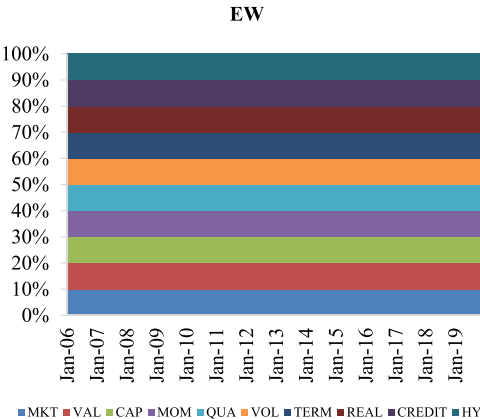
To assess whether we effectively balance the asset class and macroeconomic sensitivities of the different factors (shown in Figure 3) by combining them in a factor portfolio, we compare

²⁰As another alternative benchmark, we use a traditional 60/40 portfolio, rebalanced to a constant 60% (40%) allocation to MKT (TERM), as our benchmark. The results remain similar to those obtained with the equity market portfolio as the benchmark and are available upon request.

²¹As pointed out by Ung and Kang (2015), the cost of directly replicating a factor-based portfolio depends on the market and factors in question and can vary between 10 bps and 50 bps each way. Similar results are reported by Novy-Marx and Velikov (2016). Therefore, we regard turnover-dependent transaction costs of 50 bps as reasonably conservative.

²²Koedijk et al. (2016b) provide similar conclusions regarding the relative return and risk characteristics of equally weighted and inverse-volatility-weighted factor portfolios.

(a) Naive allocation



(b) Risk-based strategies

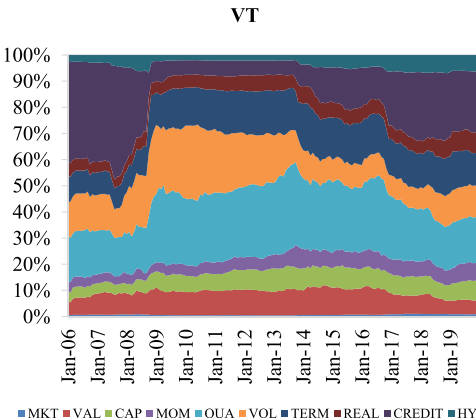
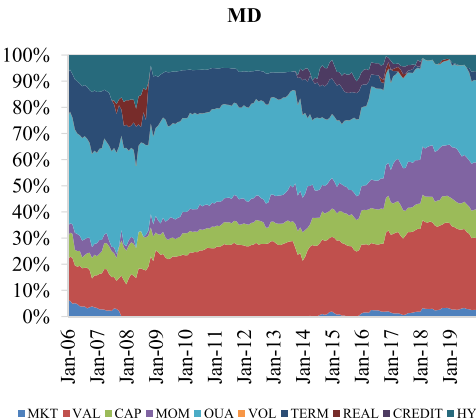
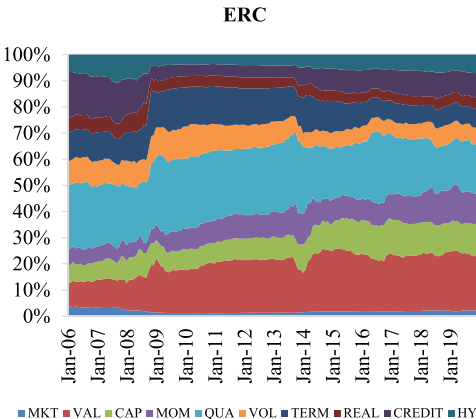
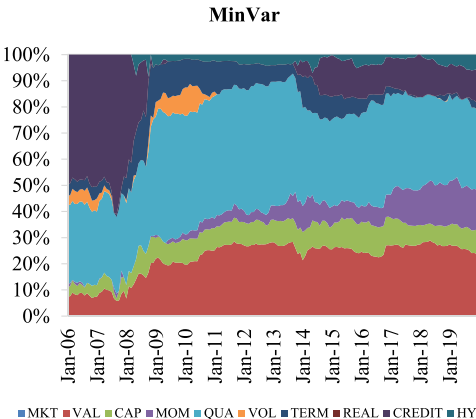


FIGURE 4 This figure depicts the weights for each “combined” factor using different factor allocation strategies (refer to Table 2 and Section 2.2) over the evaluation period from January 2006 to December 2019. The factor abbreviations are explained in Table 1

(c) Risk-and-return-based strategies

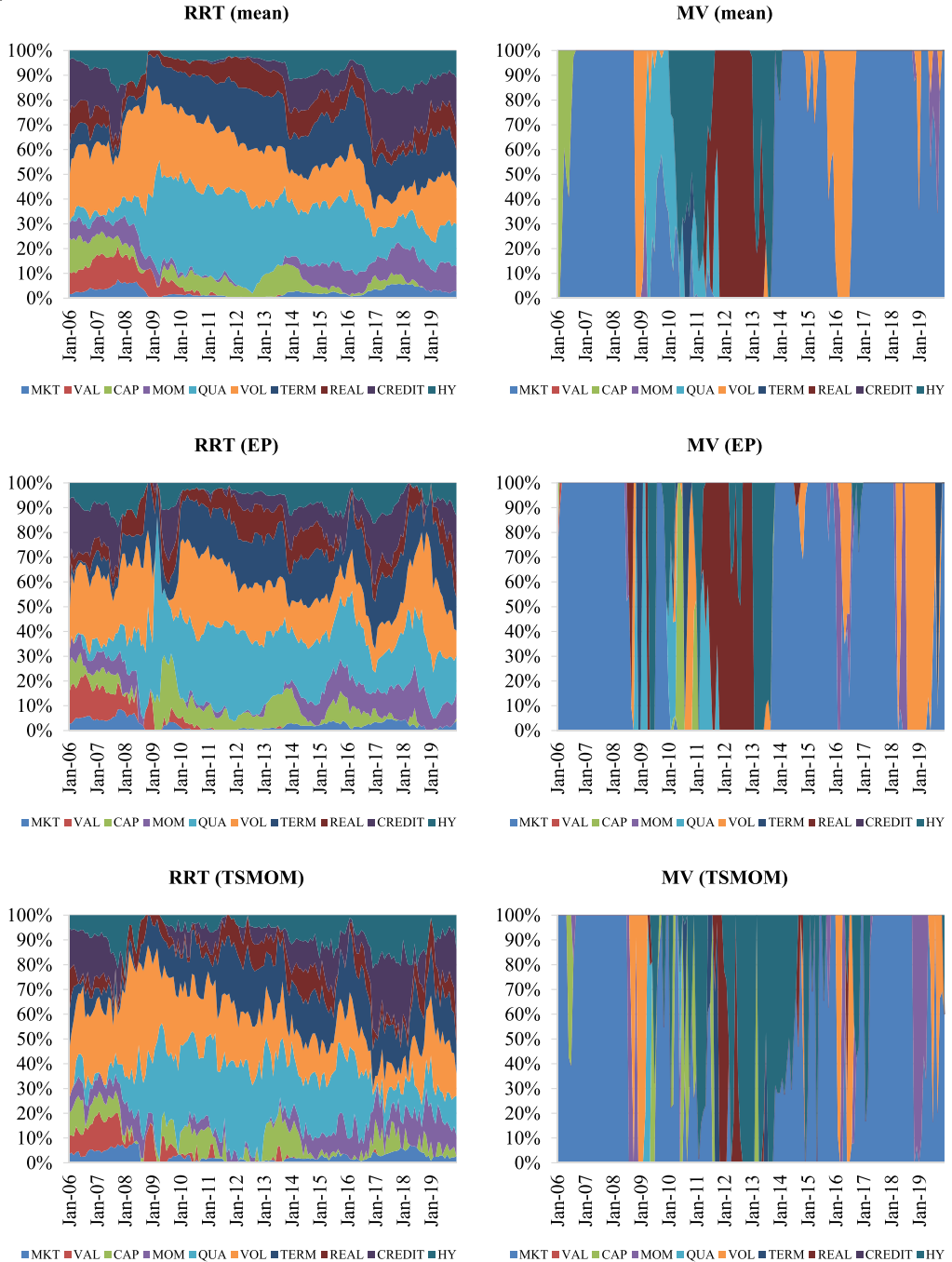


FIGURE 4 (Continued)

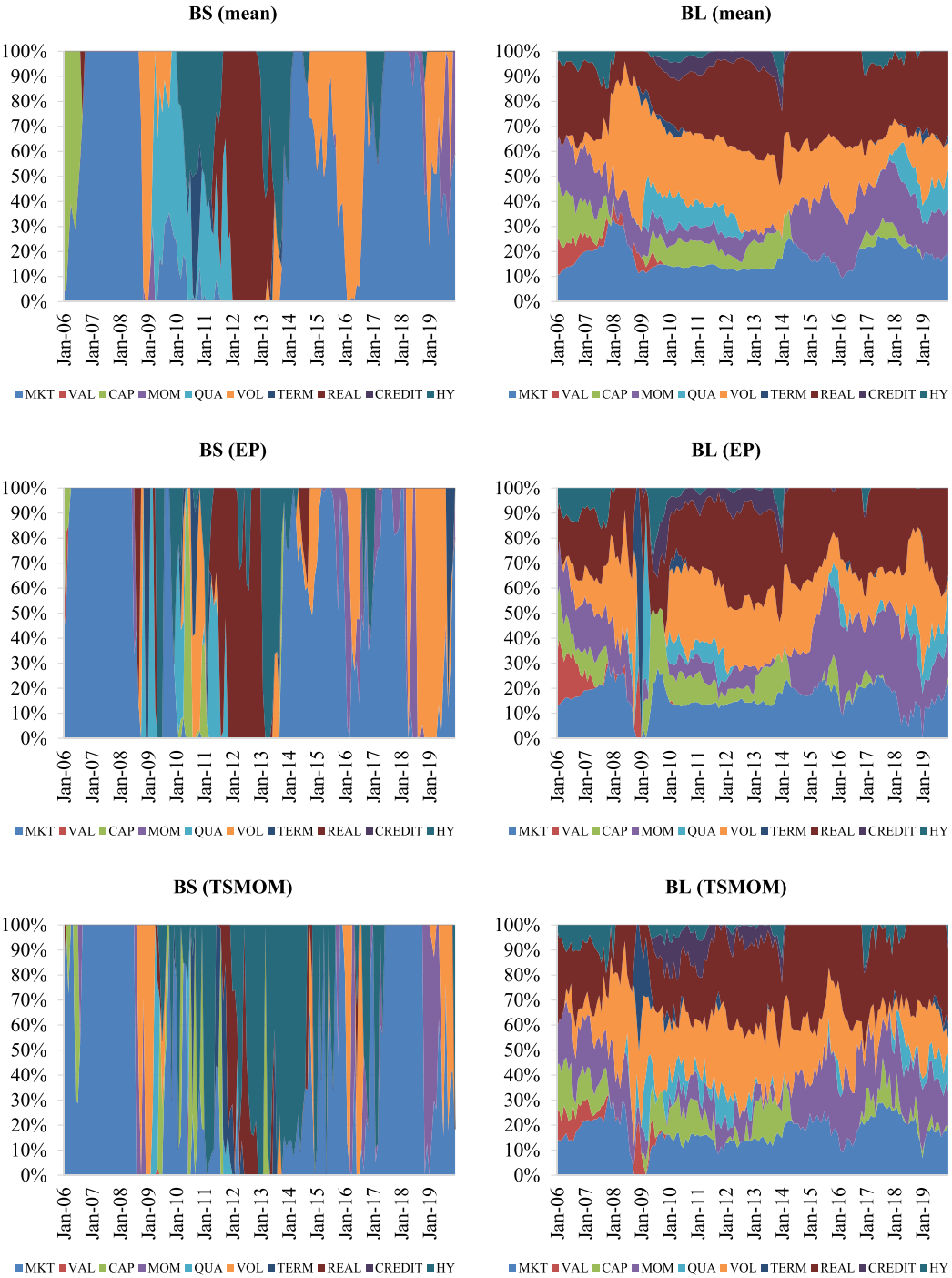
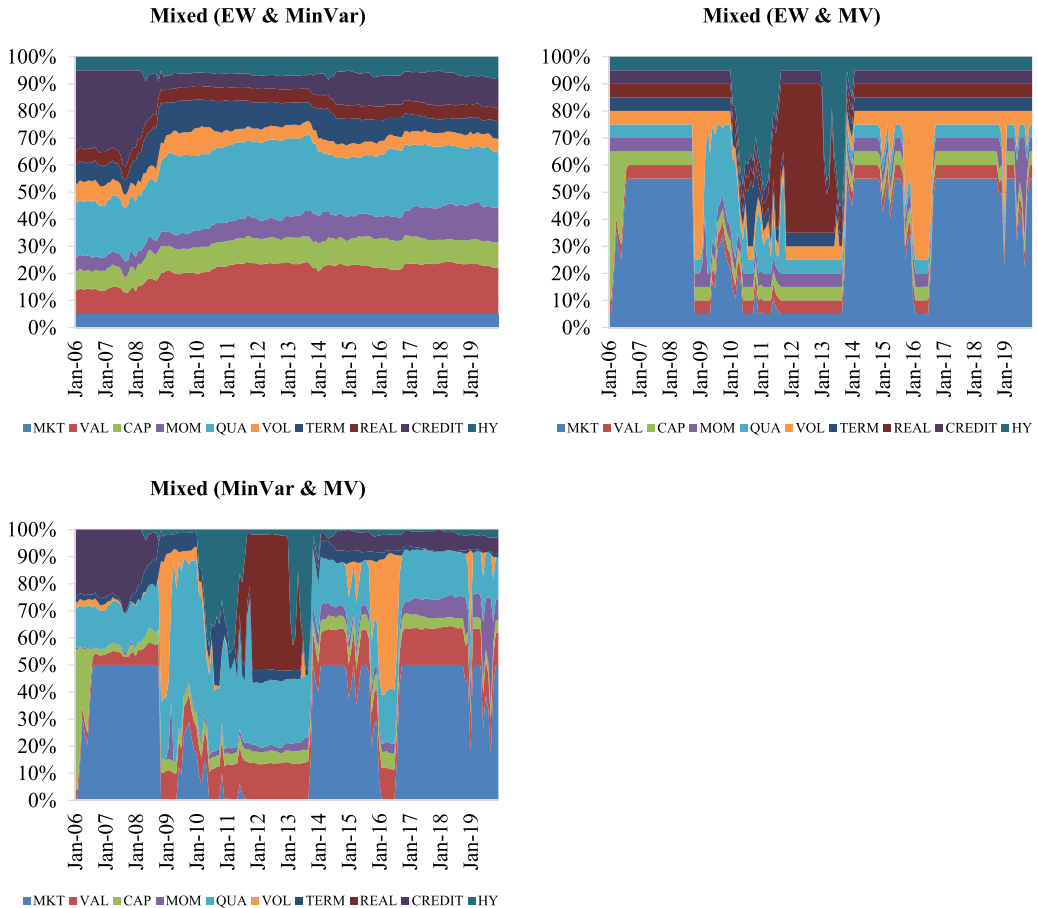


FIGURE 4 (Continued)

(d) Mixed strategies

**FIGURE 4** (Continued)

the average portfolio returns in different “states” of the world. In line with our findings so far, Figure 5 highlights that most strategies deliver positive average portfolio returns across the varying market environments. The unstable performance of the MV and BS strategies is attributable to those strategies not being well diversified (see Table 4) and heavily exposed to the equity premium with concentrated portfolio allocations (see Figure 4). This finding is in line with Blitz, Huij, Lansdorp, and van Vliet (2014), who show that if investors use a market beta overlay, that is, a 100% position in the market portfolio accompanied by long-short positions in factor portfolios, most of the risk budget is consumed by the equity premium, diminishing the diversification benefits of other factors.

Overall, our results indicate that it is possible to achieve efficient diversification across factors, but the diversification benefits are sensitive to estimation errors in expected returns (as in the case of the MV and BS strategies). They also support the view, vindicated by Asness (2016) and Asness et al. (2017), that any timing of factors does not greatly improve upon their strategic diversification. Therefore, they suggest sticking to a diversified portfolio of uncorrelated factors in which an investor believes in the long run instead of seeking to tactically time them. The stable performance of factor-based portfolios across varying

TABLE 4 Performance statistics

This table reports the performance statistics of a cash investment, the equity market portfolio (MKT), and total return portfolios consisting of a 100% position in cash with a zero-dollar position in a factor portfolio net of turnover-dependent transaction costs of 50 basis points over the evaluation period from January 2006 to December 2019: annualized arithmetic mean of monthly returns (Mean), annualized standard deviation (SD), Sharpe ratio (SR), maximum drawdown (MDD), average portfolio turnover (TO), and average diversification ratio (DR). MDD is defined as the largest cumulative loss from peak to trough, where the trough is the lowest value between two consecutive peaks; TO is defined as the average absolute change of factor portfolio weights w_{it} over the evaluation period across all N factors; and DR is defined as the average ratio of the weighted average of volatilities to total portfolio volatility (see Eq. (3)). The factor allocation strategies are described in Table 2.

| Strategy | Panel A: Single factors | | | | | | Panel B: Combined factors | | | | | |
|---------------------------------------|-------------------------|-------|------|--------|-------|-----|---------------------------|-------|-------|--------|-------|-----|
| | Mean | SD | SR | MDD | TO | DR | Mean | SD | SR | MDD | TO | DR |
| Cash | 1.2% | 0.5% | – | – | – | – | 1.2% | 0.5% | – | 0.0% | – | – |
| MKT | 7.8% | 16.5% | 0.12 | –56.1% | – | – | 7.8% | 16.5% | 0.12 | –56.1% | – | – |
| (a) Naïve allocation | | | | | | | | | | | | |
| EW | 3.4% | 3.0% | 0.76 | –10.7% | 2.0% | 2.4 | 3.5% | 3.0% | 0.79 | –10.0% | 1.8% | 2.0 |
| (b) Risk-based strategies | | | | | | | | | | | | |
| MinVar | 1.5% | 1.2% | 0.24 | –2.2% | 10.3% | 5.3 | 1.7% | 1.4% | 0.38 | –3.8% | 5.9% | 3.5 |
| ERC | 2.2% | 1.6% | 0.64 | –4.0% | 7.2% | 4.1 | 2.4% | 1.6% | 0.80 | –4.0% | 3.4% | 3.2 |
| MD | 1.6% | 1.3% | 0.32 | –2.2% | 11.4% | 5.7 | 2.2% | 1.4% | 0.76 | –2.5% | 5.9% | 3.7 |
| VT | 2.8% | 1.8% | 0.90 | –4.5% | 3.6% | 2.9 | 2.7% | 1.8% | 0.84 | –4.6% | 3.6% | 2.5 |
| (c) Risk- and return-based strategies | | | | | | | | | | | | |
| RRT (mean) | 2.7% | 2.3% | 0.68 | –6.8% | 9.5% | 2.3 | 2.6% | 2.2% | 0.65 | –6.2% | 8.8% | 2.0 |
| RRT (EP) | 2.8% | 2.2% | 0.75 | –4.7% | 17.3% | 2.3 | 2.7% | 2.2% | 0.70 | –4.5% | 15.9% | 2.0 |
| RRT (TSMOM) | 1.9% | 2.1% | 0.33 | –6.4% | 24.7% | 2.3 | 1.8% | 2.2% | 0.27 | –6.2% | 23.3% | 2.0 |
| MV (mean) | 7.4% | 16.3% | 0.39 | –59.0% | 15.8% | 1.1 | 0.4% | 11.4% | –0.07 | –50.6% | 25.6% | 1.1 |
| MV (EP) | 7.9% | 14.0% | 0.48 | –39.8% | 31.4% | 1.0 | 4.0% | 9.0% | 0.32 | –29.6% | 41.7% | 1.1 |
| MV (TSMOM) | 6.7% | 13.2% | 0.42 | –33.9% | 52.2% | 1.1 | –0.4% | 9.7% | –0.16 | –45.5% | 63.8% | 1.1 |
| BS (mean) | 5.3% | 12.6% | 0.33 | –48.2% | 23.4% | 1.2 | 0.5% | 9.8% | –0.07 | –47.0% | 27.8% | 1.2 |
| BS (EP) | 6.8% | 11.7% | 0.49 | –35.2% | 37.2% | 1.1 | 3.2% | 8.5% | 0.24 | –30.1% | 47.2% | 1.1 |
| BS (TSMOM) | 5.1% | 11.9% | 0.33 | –34.8% | 62.4% | 1.2 | –1.0% | 8.9% | –0.25 | –45.8% | 69.4% | 1.1 |
| BL (mean) | 2.4% | 5.3% | 0.23 | –25.4% | 19.4% | 1.6 | 3.3% | 4.0% | 0.52 | –17.3% | 10.6% | 1.6 |
| BL (EP) | 2.7% | 5.4% | 0.28 | –21.8% | 32.2% | 1.6 | 3.5% | 3.9% | 0.58 | –13.7% | 19.9% | 1.6 |
| BL (TSMOM) | 1.2% | 5.1% | 0.01 | –20.8% | 38.4% | 1.6 | 2.6% | 3.8% | 0.37 | –13.3% | 24.9% | 1.6 |
| (d) Mixed strategies | | | | | | | | | | | | |
| EW & MinVar | 2.5% | 1.9% | 0.68 | –5.7% | 5.9% | 3.4 | 2.6% | 2.0% | 0.74 | –6.4% | 3.5% | 2.7 |
| EW & MV | 5.4% | 9.4% | 0.45 | –38.3% | 9.6% | 1.3 | 2.0% | 6.8% | 0.12 | –31.5% | 13.8% | 1.4 |
| MinVar & MV | 4.4% | 8.4% | 0.39 | –34.4% | 13.5% | 1.4 | 1.1% | 6.0% | –0.02 | –29.2% | 15.7% | 1.4 |

market environments is worth noting, especially for institutional investors with liabilities. However, we leave the question how the presence of liabilities might already affect the initial allocation of factors (see Elton & Gruber, 1992, in the context of asset-based allocation) for future research.

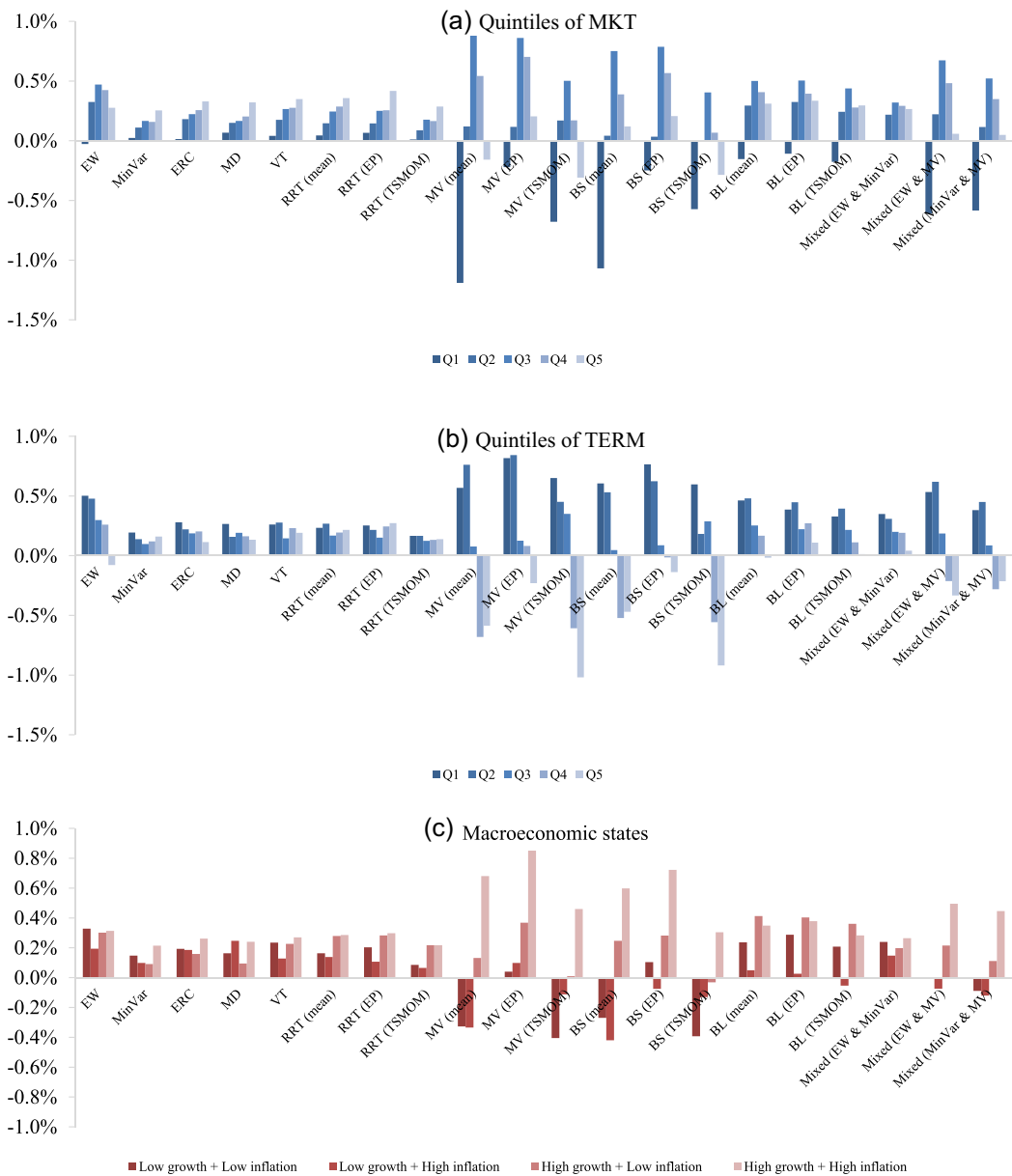


FIGURE 5 This figure depicts the average monthly portfolio returns (based on “combined” factors) over the evaluation period from January 2006 to December 2019 over quintiles of the equity premium (MKT) in panel A, over quintiles of the term premium (TERM) in panel B, and over four different macroeconomic states in panel C. In panels A and B, Q1 (Q5) denotes the quintile with the lowest (highest) premium. In panel C, a period of low (high) growth is characterized by the Chicago Fed National Activity Index (CFNAI) below (above) its sample median, and a period of low (high) inflation is characterized by the year-on-year inflation rate (CPIYOY) below (above) its sample median. The factor allocation strategies are described in Table 2

4.2 | Comparison of factor-based total return portfolios within a multiple testing framework

While our results so far suggest that many factor-based total return portfolios exhibit a better performance than a pure cash investment, and that the simple EW portfolio fares comparatively well, these analyses do not draw any statistical inference. Ultimately, we are interested in two related questions. First, does at least one total return portfolio statistically significantly outperform a cash or equity market benchmark? Second, does it pay to incorporate information on expected risk (and returns), that is, does the allocation strategy matter such that there is some total return portfolio that performs better than the one consisting of a 100% position in cash together with a zero-dollar position in an equally weighted factor portfolio? Answering these two questions requires the comparison of all total return portfolios against the pure cash or equity market investment, and of all alternative total return portfolios against the EW portfolio.

In order to carefully draw these comparisons, we revert to Hansen's (2005) SPA test, which allows us not only to compare the performance against a benchmark, but also to examine how confident we can be that the "best" portfolio among a set of multiple portfolios is genuinely better than the benchmark, given that it has been selected from a potentially large pool of portfolios. By doing so, we avoid a common bias in statistical inference referred to as "data snooping", that is, the fact that when many portfolios are evaluated individually on the same data set, some are bound to show superior performance by chance alone, even though they are not genuinely superior (Sullivan, Timmermann, & White, 1999, 2001).

In particular, we test the null hypothesis that a chosen benchmark is not inferior to any alternative total return portfolio,

$$H_0: \max_{j=1, \dots, J} E(d_{j,t}) \equiv \mu_j \leq 0$$

where $d_{j,t}$ is the difference between the performance measure of portfolio j and that of the benchmark at time t . If the null hypothesis can be rejected, there is at least one portfolio that outperforms the benchmark.

The SPA test employs the studentized test statistic

$$V_T^{\text{SPA}} = \max \left(\max_{j=1, \dots, J} \frac{\sqrt{T} \bar{d}_j}{\hat{\omega}_j}, 0 \right),$$

where $\bar{d}_j = T^{-1} \sum_{t=1}^T d_{j,t}$ denotes the average relative performance of portfolio j , and $\hat{\omega}_j^2$ is some consistent estimator for $\omega_j^2 \equiv \text{var}(\sqrt{T} \bar{d}_j)$. We obtain the distribution of the test statistic under the null by implementing the stationary bootstrap simulation approach of Politis and Romano (1994)²³ with $B = 10,000$ resamples that we recenter according to Hansen (2005). We then compute the studentized test statistic under the bootstrap $V_{b,T}^{\text{SPA}*}$ for each resample, and ultimately provide a consistent estimate of the p -value as

²³The chosen bootstrap approach involves combining blocks with random lengths that are chosen to be geometrically distributed with a mean block length of q^{-1} . In our empirical application, we set the smoothing parameter $q = 0.5$.

$$\hat{p}_{\text{SPA}} = \sum_{b=1}^B \frac{1_{\{V_{b,T}^{\text{SPA}^*} > V_T^{\text{SPA}}\}}}{B},$$

such that the null hypothesis of the SPA test is rejected for small p -values.²⁴

While the SPA test can answer the question whether there is at least one superior portfolio, it is not able to identify all such portfolios. Therefore, we implement a stepwise extension of the SPA test (step-SPA test) developed by Hsu, Hsu, and Kuan (2010) to identify all significant portfolios, while controlling the familywise error rate (FWER), that is, the probability of at least one false rejection at the pre-specified error rate $\alpha = 5\%$.

Moreover, as controlling the FWER is fairly conservative, we can increase test power by alternatively controlling the false rejections as a proportion of the number of total rejections, that is, the false discovery proportion (FDP). By implementing such an FDP-SPA test, we tolerate a higher number of false rejections, but are better able to identify superior factor portfolios. In our empirical analysis, we follow Hsu, Kuan, and Yen (2014) and choose to asymptotically control the probability of the FDP exceeding a proportion $\gamma = .1$ at the level $\alpha = 5\%$, that is, we require less than 10% of the rejected portfolios to be falsely identified.²⁵

Table 5 summarizes the results of these data-snooping tests: columns (1) and (2) indicate the chosen benchmark, that is, a cash investment (Cash), the equity market portfolio (MKT), or a 100% position in cash together with a zero-dollar position in an equally weighted factor portfolio (EW), as well as the performance measures used: mean monthly returns (Mean) and Sharpe ratios (SR). Column (3) specifies the “most significant” portfolio, that is, the portfolio with the lowest nominal p -value, which results from pairwise comparisons with the benchmark and thus does not account for the entire set of portfolios. Most importantly, columns (5) and (6) report the consistent p -value together with the lower and upper bounds as well as the number of significant portfolios identified by the step-SPA test and the FDP-SPA test, respectively.

First, the results in Table 5 confirm that, regardless of whether allocating to “single” factors (Panel A) or to “combined” factors (Panel B), there are total return portfolios that can statistically significantly outperform a pure cash investment both in terms of mean monthly total returns and Sharpe ratios when considered in isolation, as indicated by the nominal p -values below .05. Moreover, we can even reject the null hypothesis of the SPA tests at the 5% level of significance (with consistent p -values below .05) when comparing the total return portfolios against the pure cash investment, that is, the outperformance is robust to data snooping concerns. Ultimately, when allocating to “single” factors, the step-SPA and FDP-SPA tests identify the EW, VT, RRT (mean), RRT (EP), and the mixed (EW & MinVar) (only in terms of mean monthly returns) total return portfolios as superior relative to the pure cash investment. When allocating to “combined” factors, both tests additionally identify the ERC and MD total return portfolios as superior relative to cash.

Second, the results show that the equity market portfolio can be significantly outperformed by an equally weighted factor portfolio (EW) in terms of Sharpe ratios, as indicated by the consistent p -values below .01. Therefore, we conclude that diversification across factors beyond genuine risk premiums, such as the equity premium, is worthwhile at least on a risk-adjusted return basis.

²⁴Hansen (2005) shows that both an upper and a lower bound for the p -value can be obtained, whereby the upper bound assumes that all competing strategies are exactly as good as the benchmark strategy, and the lower bound assumes that the strategies with worse performance than the benchmark strategy are poor models in the limit.

²⁵See Hsu et al. (2010, 2014) for further details on the implementation of the step-SPA and FDP-SPA tests.

TABLE 5 SPA tests

This table reports the results of the SPA, step-SPA and FDP-SPA tests (with the false discovery proportion (FDP) not exceeding $\gamma = 10\%$) with the pre-specified error rate $\alpha = 5\%$ for the performance of total return portfolios consisting of a 100% position in cash together with a zero-dollar position in a factor portfolio (refer to Table 2 and Section 2.2) compared to a benchmark over the evaluation period from January 2006 to December 2019. Column (1) specifies the benchmark, that is, a cash investment (Cash), the equity market portfolio (MKT), or a total return portfolio consisting of a 100% position in cash with a zero-dollar position in an equally weighted factor portfolio (EW). Column (2) indicates the performance measure used for the data-snooping tests: mean monthly returns (Mean) and Sharpe ratios (SR). The “most significant” portfolio in column (3) is the portfolio with the lowest nominal p -value. The nominal p -value ignores the search over all portfolios that preceded the selection of the portfolio being compared to the benchmark, that is, it does not account for the entire set of portfolios. In addition, we report the consistent p -value of the SPA test, as well as its lower and upper bounds. Finally, we state the number of outperforming portfolios identified by the step-SPA and FDP-SPA tests. “-” indicates that there is no outperforming strategy at the 5% level of significance.

| Benchmark | Performance measure | Most significant portfolio | Nominal p -value | SPA p -value [lower bound, upper bound] | # of significant strategies identified by [step-SPA, FDP-SPA] |
|---------------------------|---------------------|----------------------------|--------------------|---|---|
| Panel A: Single factors | | | | | |
| Cash | Mean | RRT (mean) | .0004 | .0061 [.0061, .0061] | [5, 5] |
| | SR | VT | .0005 | .0070 [.0070, .0070] | [4, 4] |
| MKT | Mean | MV (TSMOM) | .4804 | .7524 [.6298, .7524] | [-, -] |
| | SR | EW | .0011 | .0096 [.0074, .0096] | [1, 1] |
| EW | Mean | MV (TSMOM) | .0932 | .3385 [.2046, .3848] | [-, -] |
| | SR | VT | .2377 | .6371 [.3558, .7084] | [-, -] |
| Panel B: Combined factors | | | | | |
| Cash | Mean | VT | .0025 | .0099 [.0088, .0099] | [7, 7] |
| | SR | MD | .0027 | .0100 [.0088, .0100] | [7, 7] |
| MKT | Mean | MV (TSMOM) | .8400 | .9075 [.5694, .9313] | [-, -] |
| | SR | EW | .0009 | .0085 [.0070, .0093] | [1, 1] |
| EW | Mean | MV (TSMOM) | .4176 | .8703 [.6209, .9013] | [-, -] |
| | SR | MD | .3887 | .7836 [.5786, .8527] | [-, -] |

Third, turning to the comparison of those total return portfolios that require estimates of factor moments relative to the EW portfolio, the results in Table 5 indicate that the allocation approach does not matter much, that is, no alternative factor allocation is able to deliver a statistically significant outperformance relative to the EW portfolio even in isolation (nominal p -values above .05). This finding is in line with De Miguel et al. (2009) and Hsu et al. (2018), who document that an equally weighted asset portfolio cannot be outperformed by alternative asset-based allocation strategies since the benefits of “optimal” diversification are more than offset by estimation errors. Therefore, our results contribute to the literature on optimization methods by corroborating these previous findings using a novel data set, that is, global equity and fixed income factor premia.

Taken together, our multiple testing results confirm prior research, which claims that factor investing is profitable beyond exploiting genuine risk premia. Moreover, given that the performance of the majority of factor-based portfolios is robust across various economic scenarios, our results ascertain that portfolios can be efficiently diversified using factor-based allocation. Ultimately, our analysis provides guidance for investors interested in factor-based allocation by showing that there is no strategy that can significantly outperform the EW portfolio. An investor can already reap the benefits of factor-based allocation by equally weighting the available factors, without the need to resort to more complicated strategies. Our findings might help investors and asset managers to better understand and explain the advantages of the factor methodology.

Finally, to gain a better understanding of the properties of the EW portfolio, we summarize its risk and return characteristics in Figure 6. Figure 6a,b plots its cumulative performance and drawdowns over the evaluation period, while Figure 6c depicts the risk contribution ratio of each “combined” factor contained in the portfolio, that is, its risk contribution relative to total portfolio volatility.

As visualized in Figure 6a,b, the EW portfolio exhibits a predominantly positive performance with small drawdowns. The maximum drawdown (−10.0%) is confined to the time of the global financial crisis when, in contrast, the equity premium lost more than half of its value. Figure 6c highlights that its total volatility, and thus its observed drawdowns as well, are mainly driven by the MKT and HY premia, while the VAL premium has a mitigating effect.

5 | ROBUSTNESS CHECKS

5.1 | Alternative estimation windows

The “optimal” solution of all factor-based allocation strategies (except the equally weighted factor portfolio) is sensitive to the estimates of the expected (co)variances and returns. In order to verify whether our results are robust to the choice of estimation window, we repeat the analyses in Table 5 with shorter estimation windows of 36 and 48 months, respectively. Table 6 summarizes the results.

The results are comparable to our previous findings. In each case, we identify at least one total return portfolio that significantly outperforms a pure cash investment or the equity market portfolio (at least in terms of Share ratios). Moreover, except for the case in which we compare strategies allocating to “single” factors (Panel A) based on an estimation window of 36 months using mean monthly returns as the performance measure,²⁶

²⁶The strategy identified as statistically significant by the step-SPA and FDP-SPA tests in this case is MV (EP).

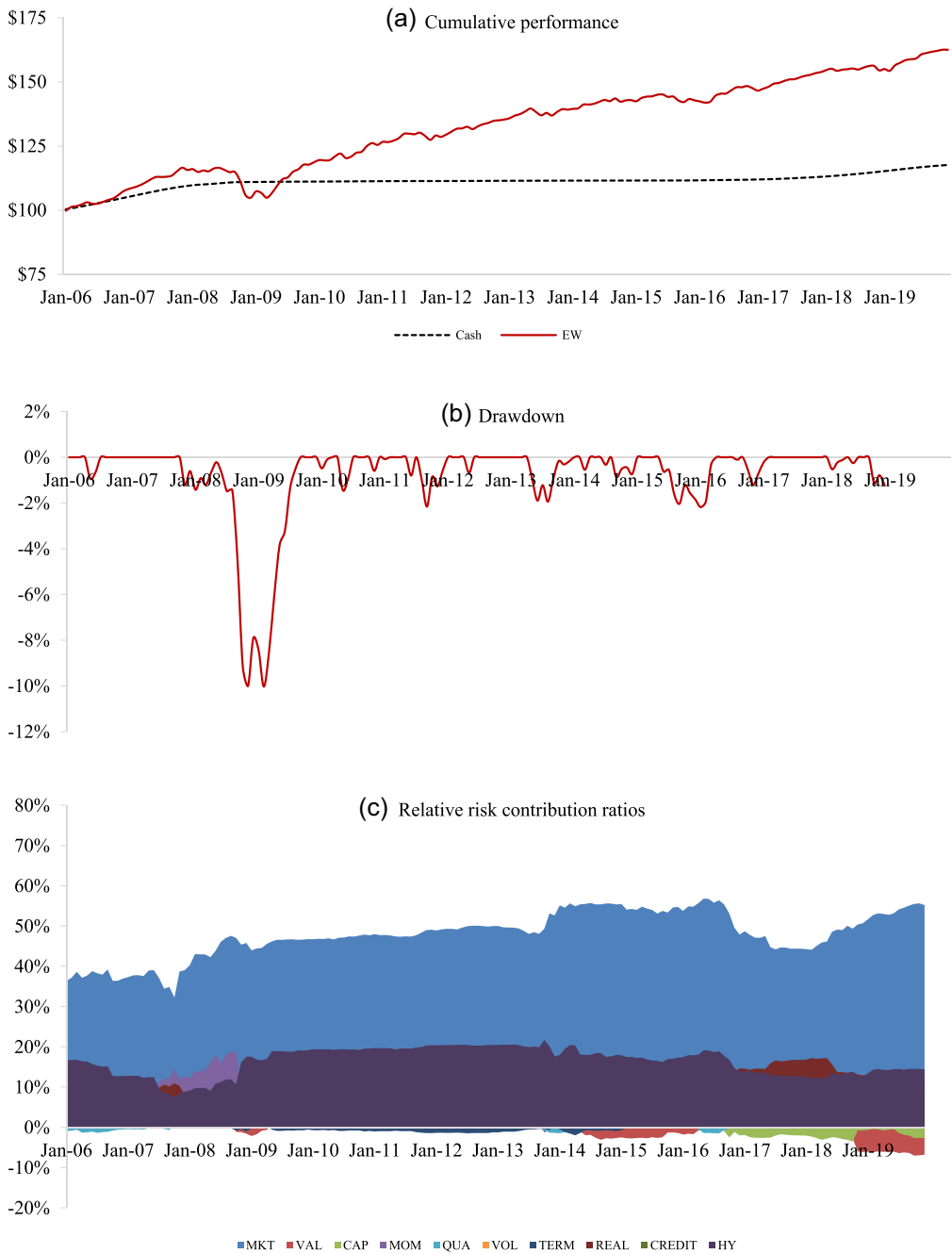


FIGURE 6 Panel A depicts the cumulative performance of a pure cash investment (Cash) and of a total return portfolio consisting of a 100% position in cash with a zero-dollar position in an equally-weighted factor portfolio based on “combined” factors (EW) net of turnover-dependent transaction costs of 50 basis points over the evaluation period from January 2006 to December 2019. Panel B plots the drawdown associated with the EW portfolio, while panel C shows the risk contribution ratio of each “combined” factor within the EW portfolio. The relative risk contribution ratio of factor i is defined as the risk contribution of factor i divided by the total portfolio volatility. The factor abbreviations are explained in Table 1

TABLE 6 Robustness check: Alternative estimation windows

This table reports the number of total return portfolios, consisting of a 100% position in cash together with a zero-dollar position in a factor portfolio using alternative estimation windows of 36 or 48 months for the expected covariance matrices and returns (refer to Table 2 and Section 2.2), that outperform a given benchmark over the evaluation period from January 2006 to December 2019 as identified by the step-SPA and FDP-SPA tests (with the false discovery proportion (FDP) not exceeding $\gamma = 10\%$) for the pre-specified error rate $\alpha = 5\%$. Column (1) specifies the benchmark, that is, a cash investment (Cash), the equity market portfolio (MKT), or a total return portfolio consisting of a 100% position in cash with a zero-dollar position in an equally weighted factor portfolio (EW). Column (2) indicates the performance measure used for the data-snooping tests: mean monthly returns (Mean) and Sharpe ratios (SR). “–” indicates that there is no outperforming strategy at the 5% level of significance.

| | | # of significant strategies identified by [step-SPA, FDP-SPA] | |
|---------------------------|---------------------|--|-----------|
| Benchmark | Performance measure | 36 months | 48 months |
| Panel A: Single factors | | | |
| Cash | Mean | [5, 5] | [5, 5] |
| | SR | [6, 6] | [5, 5] |
| MKT | Mean | [−, −] | [−, −] |
| | SR | [1, 1] | [1, 1] |
| EW | Mean | [1, 1] | [−, −] |
| | SR | [−, −] | [−, −] |
| Panel B: Combined factors | | | |
| Cash | Mean | [7, 7] | [6, 6] |
| | SR | [7, 7] | [6, 6] |
| MKT | Mean | [−, −] | [−, −] |
| | SR | [1, 1] | [1, 1] |
| EW | Mean | [−, −] | [−, −] |
| | SR | [−, −] | [−, −] |

we find no alternative factor allocation outperforming the naïve equally weighted factor portfolio.

5.2 | Alternative evaluation periods

While the bootstrap procedure implemented for the SPA tests provides random sub-samples, it still relies on the original data series and is therefore influenced by the choice of the evaluation period. Besides, our findings in Section 4.1 suggest that the average performance, especially of those strategies that rely on expected return estimates, was overly hurt by the global financial crisis. In order to verify whether our results are robust to the choice of evaluation period, we repeat our analysis in Table 5 for an extended evaluation period from January 2000 to December 2019,²⁷ and for the more recent evaluation period from January 2012 to December 2019.

²⁷Depending on data availability (see Table 1), we expand the factor set for the extended evaluation period when sufficient data become available.

TABLE 7 Robustness check: Alternative evaluation periods

This table reports the number of total return portfolios, consisting of a 100% position in cash together with a zero-dollar position in a factor portfolio for the expected covariance matrices and returns (refer to Table 2 and Section 2.2), that outperform a given benchmark over the alternative evaluation periods from January 2000 to December 2019 or from January 2012 to December 2019 as identified by the step-SPA and FDP-SPA tests (with the false discovery proportion (FDP) not exceeding $\gamma = 10\%$) for the pre-specified error rate $\alpha = 5\%$. Column (1) specifies the benchmark, that is, a cash investment (Cash), the equity market portfolio (MKT), or a total return portfolio consisting of a 100% position in cash with a zero-dollar position in an equally weighted factor portfolio (EW). Column (2) indicates the performance measure used for the data-snooping tests: mean monthly returns (Mean) and Sharpe ratios (SR). “–” indicates that there is no outperforming strategy at the 5% level of significance.

| | | # of significant strategies identified by [step-SPA, FDP-SPA] | |
|---------------------------|---------------------|--|-----------|
| Benchmark | Performance measure | 2000–2019 | 2012–2019 |
| Panel A: Single factors | | | |
| Cash | Mean | [7, 7] | [4, 4] |
| | SR | [7, 7] | [4, 4] |
| MKT | Mean | [–, –] | [–, –] |
| | SR | [3, 3] | [–, –] |
| EW | Mean | [–, –] | [–, –] |
| | SR | [–, –] | [–, –] |
| Panel B: Combined factors | | | |
| Cash | Mean | [10, 10] | [1, 1] |
| | SR | [10, 10] | [1, 1] |
| MKT | Mean | [–, –] | [–, –] |
| | SR | [4, 4] | [–, –] |
| EW | Mean | [–, –] | [–, –] |
| | SR | [–, –] | [–, –] |

The former period contains two full stock market cycles (including the dot.com bubble and the global financial crisis of 2007–2008), while the latter is a strong bull market period. Table 7 summarizes the results.

Over the extended evaluation period from January 2000 to December 2019, the performance of the total return portfolios relative to the cash investment is even stronger, with seven (10) significant strategies when allocating to single (combined) factors. Similar results are evident relative to the MKT portfolio. While the MD portfolio is identified as the most significant strategy relative to cash,²⁸ the EW portfolio is the most significant strategy relative to the MKT portfolio (in terms of Sharpe ratios). Moreover, we still cannot find evidence of any alternative factor allocation significantly outperforming the EW strategy.

For the most recent evaluation period from January 2012 to December 2019, we identify at least one strategy, the EW total return portfolio, which delivers a statistically robust

²⁸When allocating to “single” factors, the step-SPA and FDP-SPA tests additionally identify EW, MD, VT, RRT (mean), RRT (EP), and mixed (EW & MinVar) as superior to cash. When allocating to “combined” factors, MinVar, BS (mean), and BS (EP) are identified as superior as well.

outperformance relative to cash,²⁹ but no strategy significantly outperforming the MKT portfolio. Most importantly, we note that the outperformance of the EW portfolio in terms of Sharpe ratio vanishes during this short sub-sample. This finding reinforces the notion that it takes a longer time horizon to benefit from the diversification offered in factor portfolios, and our base-case sample period (used in Table 5) seems to be long enough for this effect to become observable.

6 | CONCLUSION

Factor-based allocation, that is, the deliberate diversification across factors in a portfolio context, has become increasingly popular with institutional investors. However, the question how to optimally combine factors in a portfolio has remained largely unanswered. Our study tackles this question by jointly examining the performance of factor-based allocation strategies over the evaluation period from January 2006 to December 2019. While some factor-based portfolios exhibit comparatively good economic performance, this result might be simply due to luck, given that all portfolio strategies are tested on the same data set. To control for this possibility, we apply a multiple testing framework when drawing comparisons.

Our results contribute to the growing literature on factor investing. We establish that factor-based allocation is profitable, even when implementing the necessary multiple testing corrections. We find that factor-based allocation strategies are able to achieve efficient portfolio diversification by providing positive average returns across most market environments. Moreover, we contribute to the literature on optimization methods by applying these methods to a new data set, that is, global equity and fixed income factor premia. We confirm findings from the asset class space that a naïve equal weighting of factors cannot be outperformed by more sophisticated allocation strategies, which further corroborates the view that timing benefits do not compensate for lost diversification (Asness, 2016; Asness et al., 2017). Our findings provide guidance for investors and asset managers on the best method to choose when constructing factor-based portfolios. In sum, the allocation strategy does not matter much for the diversification benefits of factor-based portfolios.

While our study is the first to rigorously compare the performance of allocation strategies within a factor optimization approach, further research into this area is certainly desirable. Interesting research questions may include the analysis of newly proposed allocation strategies, such as the hierarchical risk parity developed by López de Prado (2016), when applied to factor-based allocation or the consideration of liabilities within a factor-based allocation framework. Until then, we recommend the naïve equal weighting of factors, which is simple and cost-efficient, to reap the benefits of factor-based allocation. In particular, as advocated by Asness et al. (2017), investors should focus on those factors they believe in over the long run (based on both empirical evidence and economic theory) and diversify across these factors and harvest them in a cost-effective way. A simple strategy that avoids timing the moments of factor returns and naïvely weights all factors equally emerges as the most appropriate approach from our analyses.

REFERENCES

- Alquist, R., Israel, R., & Moskowitz, T. (2018). Fact, fiction, and the size effect. *Journal of Portfolio Management*, 45, 34–61.

²⁹When allocating to single factors, we additionally identify the VT, BS (mean), and mixed (EW & MV) total return portfolios as superior to the cash portfolio.

- Amenc, N., Deguest, R., Goltz, F., Lodh, A., & Martellini, L. (2014). Risk allocation, factor investing and smart beta: Reconciling innovations in equity portfolio construction. EDHEC-Risk Institute, Nice, France.
- Amenc, N., Goltz, F., Lodh, A., Martellini, L., & Shirbini, E. (2012). Diversifying the diversifiers and tracking the tracking error: Outperforming cap-weighted indices with limited risk of underperformance. *Journal of Portfolio Management*, 38, 72–88.
- Ang, A. (2014). Asset management: A systematic approach to factor investing, New York: Oxford University Press.
- Ang, A., Goetzmann, W. N., & Schaefer, S. M. (2009). Evaluation of active management of the Norwegian Government Pension Fund – Global. Report to the Norwegian Government.
- Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *Journal of Finance*, 61, 259–299.
- Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2009). High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics*, 91, 1–23.
- Ang, I.-C., Jiang, S., Maloney, T., & Metchick, S. (2018). *Understanding alternative risk premia* (AQR White paper), AQR Capital Management, Greenwich, CT.
- AQR Capital Management (2018). *The illusion of active fixed income alpha* (Alternative Thinking 4Q18), AQR Capital Management, Greenwich, CT.
- Arnott, R. D., Beck, N., Kalesnik, V., & West, J. (2016). How can “smart beta” go horribly wrong? Research Affiliates Article. https://www.researchaffiliates.com/en_us/publications/articles/442_how_can_smart_beta_go_horribly_wrong.html
- Arnott, R. D., Harvey, C. R., Kalesnik, V., & Linnainmaa, J. (2019). Alice's adventures in factorland: Three plunders that plague factor investing. *Journal of Portfolio Management*, 45, 18–36.
- Arnott, R. D., Kelso, C. M., Kiscadden, S., & Macedo, R. (1989). Forecasting factor returns: An intriguing possibility. *Journal of Portfolio Management*, 16, 28–35.
- Asness, C. S. (1997). The interaction of value and momentum strategies. *Financial Analysts Journal*, 53, 29–36.
- Asness, C. S. (2016). The siren song of factor timing aka “smart beta timing” aka “style timing”. *Journal of Portfolio Management*, 42, 1–6.
- Asness, C. S., Chandra, S., Imanen, A., & Israel, R. (2017). Contrarian timing is deceptively difficult. *Journal of Portfolio Management*, 43, 72–87.
- Asness, C. S., Frazzini, A., & Pedersen, L. H. (2019). Quality minus junk. *Review of Accounting Studies*, 24, 34–112.
- Asness, C. S., Imanen, A., Israel, R., & Moskowitz, T. J. (2015). Investing with style. *Journal of Investment Management*, 13, 27–63.
- Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and momentum everywhere. *Journal of Finance*, 68, 929–985.
- Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9, 3–18.
- Bass, R., Gladstone, S., & Ang, A. (2017). Total portfolio factor, not just asset, allocation. *Journal of Portfolio Management*, 43, 38–53.
- Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *Journal of Finance*, 32, 663–682.
- Bender, J., Briand, R., Nielsen, F., & Stefek, D. (2010). Portfolio of risk premia: A new approach to diversification. *Journal of Portfolio Management*, 36, 17–25.
- Bender, J., Sun, X., Thomas, R., & Zdorovtsov, V. (2018). The promises and pitfalls of factor timing. *Journal of Portfolio Management*, 44, 79–92.
- Bender, J., & Wang, T. (2015). Extending rules-based factor portfolios to a long-short framework. *Alternative Investment Analyst Review*, 4, 12–21.
- Bessler, W., & Wolff, D. (2015). Do commodities add value in multi-asset portfolios? An out-of-sample analysis for different investment strategies. *Journal of Banking and Finance*, 60, 1–20.
- Black, F., & Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, 48, 28–43.
- Blitz, D., Huij, J., Lansdorp, S., & van Vliet, P. (2014). Factor investing: Long-only versus long-short. Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2417221
- Blyth, S., Szigety, M. C., & Xia, J. (2016). Flexible indeterminate factor-based asset allocation. *Journal of Portfolio Management*, 42, 79–93.

- Brière, M., & Szafarz, A. (2017). Factor investing: Risk premia vs. diversification benefits. Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2615703
- Campbell, J. Y. (2008). Viewpoint: Estimating the equity premium. *Canadian Journal of Economics*, 41, 1–21.
- Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies*, 21, 1509–1531.
- Chen, L., Collin-Dufresne, P., & Goldstein, R. S. (2009). On the relation between the credit spread puzzle and the equity premium puzzle. *Review of Financial Studies*, 22, 3367–3409.
- Chincoli, F., & Guidolin, M. (2017). Linear and nonlinear predictability in investment style factors: Multivariate evidence. *Journal of Asset Management*, 18, 476–509.
- Chopra, V. K., & Ziemba, W. T. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. *Journal of Portfolio Management*, 19, 6–11.
- Choueifaty, Y., & Coignard, Y. (2008). Toward maximum diversification. *Journal of Portfolio Management*, 35, 40–51.
- Choueifaty, Y., Froidure, T., & Reynier, J. (2013). Properties of the most diversified portfolio. *Journal of Investment Strategies*, 2, 1–22.
- Clarke, R., de Silva, H., & Murdock, R. (2005). A factor approach to asset allocation. *Journal of Portfolio Management*, 32, 10–18.
- Clarke, R., de Silva, H., & Thorley, S. (2013). Risk parity, maximum diversification, and minimum variance: An analytic perspective. *Journal of Portfolio Management*, 32, 39–53.
- De Miguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, 22, 1915–1953.
- Dichtl, H., Drobetz, W., Lohre, H., & Rother, C. (2020). Active factor completion strategies, forthcoming: *Journal of Portfolio Management*.
- Dichtl, H., Drobetz, W., Lohre, H., Rother, C., & Vosskamp, P. (2019). Optimal timing and tilting of equity factors. *Financial Analysts Journal*, 75, 84–102.
- Dichtl, H., Drobetz, W., Neuhierl, A., & Wendt, V.-S. (2020). Data snooping in equity premium prediction, forthcoming: *International Journal of Forecasting*. <https://doi.org/10.1016/j.ijforecast.2020.03.002>
- Dimson, E., Marsh, P., & Staunton, M. (2018). Credit Suisse global investment returns yearbook 2018, Credit Suisse AG Research Institute, Zurich, Switzerland.
- Elton, E. J., & Gruber, M. J. (1992). Optimal investment strategies with investor liabilities. *Journal of Banking and Finance*, 16, 869–890.
- Fama, E. F., & French, K. R. (2012a). Size, value, and momentum in international stock returns. *Journal of Financial Economics*, 105, 457–472.
- Fama, E. F., & French, K. R. (2012b). Volatility and premiums in US equity returns. Working paper. <https://famafrench.dimensions.com/essays/volatility-and-premiums.aspx>
- Person, W., & Harvey, C. (1993). The risk and predictability of international equity returns. *Review of Financial Studies*, 6, 527–566.
- Fleming, J., Kirby, C., & Ostdiek, B. (2001). The economic value of volatility timing. *Journal of Finance*, 56, 329–352.
- Fleming, J., Kirby, C., & Ostdiek, B. (2003). The economic value of volatility timing using “realized” volatility. *Journal of Financial Economics*, 67, 473–509.
- Fung, W., & Hsieh, D. (1997). Empirical characteristics of dynamic trading strategies: The case of hedge funds. *Review of Financial Studies*, 10, 275–302.
- Greenberg, D., Babu, A., & Ang, A. (2016). Factors to assets: Mapping factor exposures to asset allocations. *Journal of Portfolio Management*, 42, 18–27.
- Gupta, T., & Kelly, B. (2019). Factor momentum everywhere. *Journal of Portfolio Management*, 45, 1–24.
- Hamdan, R., Pavlowsky, F., Roncalli, T., & Zheng, B. (2016). A primer on alternative risk premia. Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2766850
- Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business and Economic Statistics*, 23, 365–380.
- Haugen, R. A., & Heins, A. J. (1975). Risk and the rate of return on financial assets: Some old wine in new bottles. *Journal of Financial and Quantitative Analysis*, 10, 775–784.
- Horowitz, J. L., Loughran, T., & Savin, N. E. (2000). The disappearing size effect. *Research in Economics*, 54, 83–100.

- Hsu, J., Kalesnik, V., & Viswanathan, V. (2015). A framework for assessing factors and implementing smart beta strategies. *Journal of Index Investing*, 6, 89–97.
- Hsu, P.-H., Han, Q., Wu, W., & Cao, Z. (2018). Asset allocation strategies, data snooping, and the 1/N rule. *Journal of Banking and Finance*, 97, 257–269.
- Hsu, P.-H., Hsu, Y.-C., & Kuan, C.-M. (2010). Testing the predictive ability of technical analysis using a new stepwise test without data snooping bias. *Journal of Empirical Finance*, 17, 471–484.
- Hsu, Y.-C., Kuan, C.-M., & Yen, M.-G. (2014). A generalized stepwise procedure with improved power for multiple inequalities testing. *Journal of Financial Econometrics*, 12, 730–755.
- Huang, D., Jiang, F., Tu, J., & Zhou, G. (2017). Forecasting stock returns in good and bad times: The role of market states. Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2188989
- Idzorek, T. (2007). A step-by-step guide through the Black-Litterman model: Incorporating user specified confidence levels. In Satchell S. (Ed.), *Forecasting Expected Returns in the Financial Markets*. Amsterdam: Elsevier.
- Idzorek, T. M., & Kowara, M. (2013). Factor-based asset allocation vs. asset-class-based asset allocation. *Financial Analysts Journal*, 69, 1–11.
- Ilmanen, A. (2011). *Expected returns: An investor's guide to harvesting market rewards*, Chichester, UK: John Wiley & Sons.
- Ilmanen, A., & Kizer, J. (2012). The death of diversification has been greatly exaggerated. *Journal of Portfolio Management*, 38, 15–27.
- Ilmanen, A., Maloney, T., & Ross, A. (2014). Exploring macroeconomic sensitivities: How investments respond to different economic environments. *Journal of Portfolio Management*, 40, 87–99.
- James, W., & Stein, C. (1961). Estimation with quadratic loss. In J. Neyman (Ed.), *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 361–379). Berkeley, CA: University of California Press.
- Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48, 65–91.
- Jorion, P. (1985). International portfolio diversification with estimation risk. *Journal of Business*, 58, 259–278.
- Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, 21, 279–292.
- Kan, R., & Zhou, G. (2007). Optimal portfolio choice under parameter uncertainty. *Journal of Financial and Quantitative Analysis*, 42, 621–656.
- Kirby, C., & Ostdiek, B. (2012). It's all in the timing: Simple active portfolio strategies that outperform naïve diversification. *Journal of Financial and Quantitative Analysis*, 47, 437–467.
- Koedijk, K. G., Slager, A. M. H., & Stork, P. A. (2016a). A trustee guide to factor investing. *Journal of Portfolio Management*, 42, 28–38.
- Koedijk, K. G., Slager, A. M. H., & Stork, P. A. (2016b). Investing in systematic factor premiums. *European Financial Management*, 22, 193–234.
- Kritzman, M., Page, S., & Turkington, D. (2010). In defense of optimisation: The fallacy of 1/N. *Financial Analysts Journal*, 66, 31–39.
- Lee, W. (2017). Factors timing factors. *Journal of Portfolio Management*, 43, 66–71.
- Leitch, G., & Tanner, J. E. (1991). Economic forecast evaluation: Profits versus the conventional error measures. *American Economic Review*, 81, 580–590.
- Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *Journal of Finance*, 20, 587–615.
- López de Prado, M. (2016). Building diversified portfolios that outperform out of sample. *Journal of Portfolio Management*, 42, 59–69.
- Ludvigson, S. C., & Ng, S. (2009). Macro factors in bond risk premia. *Review of Financial Studies*, 22, 5027–5067.
- Maeso, J., Martellini, L., & Rebonato, R. (2019). *Factor investing in fixed-income – cross-sectional and time-series momentum in sovereign bond markets*. EDHEC-Risk Institute.
- Maillard, S., Roncalli, T., & Teiletche, J. (2010). On the properties of equally weighted risk contributions portfolios. *Journal of Portfolio Management*, 36, 60–70.
- Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7, 77–91.
- Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8, 323–361.

- Meucci, A. (2010). The Black-Litterman approach: Original model and extensions. Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1117574
- Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, 34, 768–783.
- Naik, V., Devarajan, M., Nowobilski, A., Page, S., & Pedersen, N. (2016). *Factor investing and asset allocation: A business cycle perspective* (CFA Institute Research Foundation Series 2016-4), Charlottesville, VA: CFA Institute Research Foundation.
- Neely, C. J., Rapach, D. E., Tu, J., & Zhou, G. (2014). Forecasting the equity risk premium: The role of technical indicators. *Management Science*, 60, 1772–1791.
- Novy-Marx, R. (2014). Quality investing. Working paper, University of Rochester.
- Novy-Marx, R., & Velikov, M. (2016). A taxonomy of anomalies and their trading costs. *Review of Financial Studies*, 29, 104–147.
- Piotroski, J. D. (2000). Value investing: The use of historical financial statement information to separate winners from losers. *Journal of Accounting Research*, 38, 1–41.
- Politis, D. N., & Romano, J. P. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *Annals of Statistics*, 22, 2031–2050.
- Pukthuanthong, K., Roll, R., & Subrahmanyam, A. (2019). A protocol for factor identification. *Review of Financial Studies*, 32, 1573–1607.
- Qian, E. (2006). On the financial interpretation of risk contribution: Risk budgets do add up. *Journal of Investment Management*, 4, 41–51.
- Qian, E. (2011). Risk parity and diversification. *Journal of Investing*, 20, 119–127.
- Reinganum, M. R. (1981). Misspecification of capital asset pricing: Empirical anomalies based on earnings' yields and market values. *Journal of Financial Economics*, 9, 19–46.
- Roncalli, T. (2013). *Introduction to risk parity and budgeting*. Boca Raton, FL: Chapman & Hall/CRC Press.
- Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13, 341–360.
- Satchell, S., & Scowcroft, A. (2000). A demystification of the Black Litterman model: Managing quantitative and traditional portfolio constructions. *Journal of Asset Management*, 1, 138–150.
- Schwert, G. W. (2003). Anomalies and market efficiency. In G. M. Constantinides, M. Harris & R. Stulz (Eds.), *Handbook of the Economics of Finance*. Amsterdam: Elsevier.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19, 425–442.
- Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In J. Neyman (Ed.), *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* (1, pp. 197–206). Berkeley, CA: University of California Press.
- Sullivan, R., Timmermann, A., & White, H. (2001). Dangers of data mining: The case of calendar effects in stock returns. *Journal of Econometrics*, 105, 249–286.
- Sullivan, R., Timmermann, A., & White, H. (1999). Data-snooping, technical trading rule performance, and the bootstrap. *Journal of Finance*, 54, 1647–1691.
- Timmermann, A. (2006). Forecast combinations. In G. Elliott, C. Granger & A. Timmermann (Eds.), *Handbook of Economic Forecasting Volume 1*. Amsterdam: Elsevier.
- Tu, J., & Zhou, G. (2011). Markowitz meets Talmud: A combination of sophisticated and naïve diversification strategies. *Journal of Financial Economics*, 99, 204–215.
- Ung, D., & Kang, X. (2015). Practical considerations for factor-based asset allocation. *Journal of Index Investing*, 5, 33–47.
- Wellington, W. (2016). *A vanishing value premium?* Austin, TX, USA: Dimensional Fund Advisors LP.
- Zhang, L. (2005). The value premium. *Journal of Finance*, 60, 67–103.

How to cite this article: Dichtl H, Drobetz W, Wendt V-S. How to build a factor portfolio: Does the allocation strategy matter? *Eur Financ Manag*. 2021;27:20–58. <https://doi.org/10.1111/eufm.12264>