A Comparison of Global Factor Models

Latest version here

Matthias X. Hanauer
TUM School of Management
Technical University of Munich

First version: November 2019; This draft: March 2020

Abstract

I compare commonly employed factor models across 50 non-U.S. developed and emerging market countries by ranking them based on their maximum Sharpe ratios. Consistent with the U.S. evidence presented in Barillas, Kan, Robotti, and Shanken (2019), I find that the factor models of Fama and French (2015, 2018), Hou, Xue, and Zhang (2015), and Stambaugh and Yuan (2017) are dominated by a six-factor model that includes cash-based profitability and momentum factors, as well as a value factor that is updated monthly. The result is robust in out-of-sample tests, across subperiods, across global regions, and to methodological changes. The main problem for the dominated factors models is that they do not explain the monthly updated value factor. Hence, I conclude that the value factor is not redundant.

Keywords: Empirical asset pricing; Factor models; Value; Momentum; Profitability

JEL Classifications: G12, G14, G15

I am grateful to Patrick Bielstein, David Blitz, Jonathan Fletcher, Eric van den Heuvel, Xander Gerrmann, Wesley Gray, Corey Hoffstein, Heiko Jacobs, Raymond Kan, Sebastian Müller, Simon Rottke, Esad Smajlbegovic, Laurens Swinkels, Gertjan Verdickt, Pim van Vliet, Jack Vogel, Radovan Vojtko, Christian Walkshäusl, and Steffen Windmüller for helpful comments and suggestions. I thank Steffen Windmüller for his assistance with data. Furthermore, I am grateful for the provision of Matlab code to conduct the model comparison tests using the Sharpe measure by Cesare Robotti. Disclosures: Hanauer is also employed by Robeco, an asset management firm that, among other strategies, also offers factor investing strategies. The views expressed in this paper are those of the author and not necessarily shared by Robeco. Any remaining errors are my own.
Email: matthias.hanauer@tum.de

1 Introduction

Researchers attempt to explain differences in expected returns with factor models that use a parsimonious set of factors. The capital asset pricing model (CAPM) developed by Sharpe (1964) and Lintner (1965) was one of the earliest of these models and contains only one factor, the value-weighted market portfolio of all financial assets. Despite its theoretical appeal, the CAPM does not perform well empirically in explaining cross-sectional differences in average stock returns. Fama and French (1992) synthesize the empirical evidence up to that time (e.g., Basu, 1977; Banz, 1981; Rosenberg, Reid, and Lanstein, 1985) and conclude that if one controls for size effects, market beta does not, whereas size and book-to-market equity explain cross-sectional differences in average returns. This analysis led to the Fama and French (1993) three-factor model, which contains market, size, and value factors. This model was for many years the industry standard, sometimes augmented with the momentum factor of Jegadeesh and Titman (1993), as in Carhart (1997). Fama and French (2015) extend their three-factor model to a five-factor model by including profitability and investment factors. The models of Fama and French (1993, 2015) compete with models such as the Hou et al. (2015) q-factor model, the Stambaugh and Yuan (2017) mispricing model, and a revised six-factor model of Fama and French (2018), among others.

When comparing models, empirical studies often focus on specification tests such as the Gibbons, Ross, and Shanken (1989, GRS) test of mean-variance efficiency, in which competing models are judged based on the intercepts they leave on a broad set of test assets. Prominent examples for test assets are 25 portfolios defined by size and book-to-market sorts (see, e.g., Fama and French, 2012, 2015, 2016, 2017; Stambaugh and Yuan, 2017) or decile portfolios using various characteristics (see, e.g., Hou et al., 2015; Hou, Xue, and Zhang, 2018). However, Barillas and Shanken (2017) point out that such an approach can be problematic when it ignores the pricing impact of factors from other models among the test assets; they demonstrate that when comparing models, the only relevant comparison is how well models price the factors not included in the model, and that surprisingly, the choice of test assets is irrelevant for model comparison. Barillas and Shanken (2018) built on this observation and develop a Bayesian approach to identify a parsimonious model that produces the highest Sharpe ratio. They find that the models proposed by Fama and French (2015) and Hou et al. (2015) are dominated by models that include a momentum factor, as well as value and profitability factors that are updated monthly. Instead, of Bayesian statistics, Barillas et al. (2019) focus directly on comparing models' maximum squared Sharpe ratios using classical statistics. Although they point out that the Bayesian and classical approaches can provide complementary insights about model performance, they again conclude that a variant of the Fama and French six-factor model, that includes a cash-based profitability factor, a momentum factor and a value factor that is updated monthly, emerges as the dominant model.

An important limitation of the papers mentioned above is that all of them are based solely on United States (U.S.) market data. This observation reflects the findings of Karolyi (2016, p. 2049) that empirical research focused on non-U.S. studies are underrepresented, as only "16% (23%) of all empirical studies published in the top four (fourteen) Finance journals examine non-US markets, a fraction that is well below measures reflecting their economic importance." Drawing general conclusions solely from U.S. data can be dangerous, as research has identified cases in which results obtained from U.S. market data have proven not to hold (entirely) in international markets (e.g., Goyal and Wahal, 2015; Jacobs and Müller, 2020). As with any finding in empirical research, the results documented for the U.S. could be due to chance or data snooping and would therefore be sample-specific. To alleviate these concerns, out-of-sample evidence such as model comparisons using data from outside the U.S. is necessary. However, comparisons of a broad set of popular factor models for international markets are scarce.[1]

This study is the first to compare a broad set of common and recently proposed factor models using a comprehensive non-U.S. sample. Specifically, I investigate the following models for 50 international markets from 1990 to 2018: (i) the CAPM, (ii) the Fama and French (1993) three-factor model, (iii) the Fama and French (2015) five-factor model, (iv) the Fama and French six-factor model (the five-factor model augmented with momentum), (v) the Fama and French (2018) six-factor model that substitutes the operating profitability factor used in their original six-factor

[1]Ammann, Odoni, and Oesch (2012) and Walkshäusl and Lobe (2014) compare only the Fama and French (1993) three-factor model and a precursor of the Hou et al. (2015) q-factor model using a sample of European Monetary Union and 40 non-U.S. countries, respectively, but do not include more recent models such as the Fama and French (2015) five-factor model, the Stambaugh and Yuan (2017) mispricing model, or the six-factor model proposed in Barillas et al. (2019). Fletcher (2018, 2019) applies the Bayesian approach of Barillas and Shanken (2018) to examine the factor models of Fama and French (1993, 2015), with and without a momentum factor, using the standard value or a more timely value factor, and with the standard or with the small ends of the factors in global and U.K stock returns, respectively, but does not investigate the Hou et al. (2015) q-factor, the Stambaugh and Yuan (2017) mispricing, the Fama and French (2018) six-factor, or the six-factor model proposed in Barillas et al. (2019). Furthermore, Fletcher (2019) also conducts classical Sharpe ratio comparison tests and includes the behavioral factor model proposed by Daniel, Hirshleifer, and Sun (2019).

model with a cash-based operating profitability factor, (vi) the Hou et al. (2015) q-factor model that includes a ROE based profitability and an investment factor along with the market and the size factor, (vii) the Stambaugh and Yuan (2017) mispricing model that extends the CAPM by adding a size factor and two composite mispricing factors, management and performance, and (viii) the six-factor model proposed in Barillas et al. (2019) that substitutes the value factor used in the Fama and French (2018) six-factor model with a monthly updated value factor. The model selection follows Barillas and Shanken (2018) and Barillas et al. (2019), but I also include the CAPM and the Fama and French (1993) three-factor model to compare newer models with the older models.

I adopt a common set of construction rules for my set of global factors to improve comparability. All factors are constructed via the "traditional" 2 x 3 sorts on size and the respective factor criteria as in Fama and French (2012, 2017) for international data to ensure that differences in the factor construction procedure do not drive the results (see, e.g., also the concerns regarding deviating factor construction procedures in Hou, Mo, Xue, and Zhang, 2019). As in Walkshäusl and Lobe (2014), Jacobs (2016), and Jacobs and Müller (2020), I use annual accounting information as the reporting of quarterly accounting data is not common practice in many countries outside the U.S. at the beginning of my sample period.[2] As a consequence, I also do not include the behavioral factor model of Daniel et al. (2019) in the analysis as the model includes a post-earnings-announcement-drift factor (PEAD) that is based on abnormal returns around quarterly earnings announcement dates.

My main findings are summarized as follows. First, I document that all employed factors exhibit positive and, other than the market and size factors, significant average returns outside the U.S. Furthermore, the documented patterns are quite similar to what is reported in Barillas et al. (2019) for the U.S. Second, I find that the six-factor model proposed in Barillas et al. (2019) spans an annualized maximum Sharpe ratio of 2.36 that is substantially and significantly higher than those of competing factor models. The winning model shows a maximum Sharpe ratio improvement of

[2] This requirement is likely to have to biggest impact on the ROE based profitability factor in the Hou et al. (2015) q-factor model as a ROE factor based on the most recently announced quarterly earnings profits from the post-earnings-announcement drift as shown in Novy-Marx (2015). However, allowing the use of quarterly accounting data, could also improve the performance of the operating profitability factors in other models. Furthermore, Novy-Marx (2015) raises the concern that using quarterly accounting data introduces a look-ahead bias, because quarterly earnings include revisions and thus might be different from the earnings that are actually announced. To ensure that my results are not driven by the definition of the profitability factor, I replace the ROE based profitability factor in the Hou et al. (2015) q-factor model with a cash-based operating profitability factor in Section 5. However, my main conclusions remain the same.

about 20% and 50% compared to the Fama and French six-factor models with cash-based operating profitability and operating profitability factors, respectively. Compared to the remaining models, the improvement is even larger. Therefore, I conclude that a six-factor model that includes factors for cash-based profitability and momentum factors, and a monthly updated value factor, dominates the factor models of Fama and French (1993, 2015, 2018), Hou et al. (2015), and Stambaugh and Yuan (2017) in international markets. This finding is consistent with the U.S. evidence presented in Barillas et al. (2019). Third, I demonstrate that the main problem with the dominated factors models is that they cannot explain the returns of the monthly updated value factor. Hence, I conclude in contrast to Fama and French (2015) that the value factor is not redundant. Finally, my results are robust in out-of-sample tests, across five-year subperiods, across different regions (Asia Pacific, Europe, Japan, and emerging markets), and to various methodological changes.

This study contributes to the literature in at least three aspects. First, it adds to current research that looks for a parsimonious factor model that spans the tangency portfolio for traded factors and does not retain redundant factors. I complement the analysis in Barillas et al. (2019) for the U.S. by evaluating factor models based on their maximum Sharpe ratios in markets outside the U.S. According to Hou et al. (2018), replication makes a contribution when extending existing studies out-of-sample. Harvey (2017) further supports the replication argument by stating that many published results would not hold up under scrutiny because of unreported tests, testing of multiple hypotheses, and data snooping. Harvey, Liu, and Zhu (2016) link data snooping concerns with the strong pressures and incentives to publish, which generates a publication bias; they propose higher t-statistic hurdles. Since this is a comparison study that applies documented models to international markets, an affirmative result for this out-of-sample test alleviates concerns that the patterns found in the U.S. market are due to chance or data snooping. I use a uniform global dataset and a common set of factor construction rules, and proceed with the same statistical tests, reporting the results homogeneously across factors and factor models. In this way, I overcome potential concerns of data mining, multiple hypothesis testing, and Type I errors.

Second, I add to the discussion of the importance and validity of the value factor, and the best way to define the profitability factor. The q-factor model proposed by Hou et al. (2015) and the mispricing model in Stambaugh and Yuan (2017) do not contain a value factor. Furthermore, Fama and French (2015) find that by adding profitability and investment factors, the value factor becomes redundant. Additionally, studies such as Lev and Srivastava (2019) and Park (2019) argue that the value factor has recently failed and/or that the value definition should be adjusted. In contrast, Asness, Frazzini, Israel, and Moskowitz (2015), Barillas and Shanken (2018), and Barillas et al. (2019) demonstrate that a value factor that is updated monthly is not redundant for the U.S. I confirm this result for global markets outside the U.S. and show that dominated factors models fail primarily because they cannot explain the monthly updated value factor. I also note that while the standalone performance of the monthly updated value factor is flat over the last decade, when controlling for other factors, the monthly updated value factor shows a strong and robust performance over time. Similarly, the models proposed in Fama and French (2015, 2018), Hou et al. (2015), Stambaugh and Yuan (2017), Barillas et al. (2019) contain profitability factors with various definitions. Barillas et al. (2019, p. 14) state that "the choice of profitability factor is a key" and Ball, Gerakos, Linnainmaa, and Nikolaev (2016) show that cash-based operating profitability (a measure that excludes accruals) outperforms profitability measures that include accruals. Consistent with the results in Ball et al. (2016) and Hanauer and Huber (2019), I show that cash-based profitability measures are superior to profitability measures that include accruals in international markets.

Finally, I add to the expanding literature on international asset pricing. Previous research in this context includes studies that investigate the cross-sectional predictability of individual signals outside the U.S. For example, see, e.g., Fama and French (1998), Rouwenhorst (1998), McLean, Pontiff, and Watanabe (2009), Heston and Sadka (2010), Watanabe, Xu, Yao, and Yu (2013), and Walkshäusl and Lobe (2015), who analyze value, momentum, share issuance, seasonality, asset growth, and enterprise multiple patterns, respectively. Other studies investigate factor models in international (developed, emerging, and frontier) markets (see, for example, Griffin, 2002; Hou, Karolyi, and Kho, 2011; Fama and French, 2012, 2017; Cakici, Fabozzi, and Tan, 2013; Zaremba and Czapkiewicz, 2017; Hanauer and Lauterbach, 2019; de Groot, Pang, and Swinkels, 2012; Zaremba and Maydybura, 2019).

The results shown in this study have important implications for defining a powerful empirical asset pricing model. Since my findings confirm the evidence for the U.S. market in Barillas et al. (2019), I conclude that the six-factor model proposed by Barillas et al. (2019) robustly outperforms competing factor models for markets around the globe, and therefore it should be the preferred benchmark model for future research both for the U.S. market and globally. A powerful factor model should also be able to reduce the number of remaining anomalies in the cross-section of stock returns. For instance, Ball et al. (2016) demonstrate that a model that incorporates a cash-based profitability factor subsumes the accrual anomaly, while the Fama and French (2015) five-factor and the Hou et al. (2015) q-factor models cannot (cf., Fama and French, 2015; Hou et al., 2015). Similarly, a model that contains both a monthly updated value factor and a momentum factor raises the bar for alternative value factors. An example would be a valuation multiple that is updated monthly and shows anomalous returns relative to models that include only the standard value factor and the momentum factor. However, the anomalous returns could be due to the monthly updating frequency, while the multiple itself may not carry any unique information.

The remainder of this study is structured as followed: Section 2 describes the data, factors, and factor models. Section 3 and Section 4 present my methodology and the empirical results for the model comparison tests. Section 5 assesses the robustness of my results and Section 6 concludes.

2 Data and Factor Models

2.1 Data

My global (ex U.S.) sample comprises market data from Datastream and accounting data from Worldscope. I process the data through common static and dynamic screens to ensure data quality. As a first step, I identify stocks using Thomson Reuters Datastream's constituent lists. I use Worldscope lists, research lists, and—to eliminate survivorship bias—dead lists. Following Ince and Porter (2006), Griffin, Kelly, and Nardari (2010), and Schmidt, Von Arx, Schrimpf, Wagner, and Ziegler (2017), I apply generic as well as country-specific static screens to eliminate non-common equity stocks. In addition, I apply dynamic screens for stock returns and price data as recommended in the literature. Appendix A.1 describes the static and dynamic screens in detail. Finally, I require stocks to have market capitalization data for June of year y and the previous month to be included in months ranging from July of year y to June of year $y+1$.

The sample comprises data from 51 developed and emerging market countries. The country selection follows the Morgan Stanley Capital International (MSCI) Developed and Emerging Market Indices. I include all countries classified either as a developed or emerging market at some point during the sample period.[3] More precisely, the countries are part of the sample only in those years in which they are included in either the MSCI Developed or Emerging Markets Index. The following countries meet this criterion: Argentina, Australia, Austria, Belgium, Brazil, Canada, Chile, China, Colombia, Czech Republic, Denmark, Egypt, Finland, France, Germany, Greece, Hong Kong, Hungary, Indonesia, India, Ireland, Israel, Italy, Japan, Jordan, South Korea, Malaysia, Mexico, Morocco, the Netherlands, New Zealand, Norway, Pakistan, Peru, Philippines, Poland, Portugal, Qatar, Russia, Singapore, South Africa, Sri Lanka, Spain, Switzerland, Sweden, Thailand, Turkey, Taiwan, United Arab Emirates, the United Kingdom (U.K.), and Venezuela. The countries are only part of the final sample in those months for which at least 30 stock-month observations are available after filters. This requirement leads to the exclusion of Venezuela from the final sample.

Following Fama and French (2012, 2017), I combine developed market countries into three regions: (i) Asia Pacific (ex Japan), (ii) Europe, and (iii) Japan. Emerging markets are combined into a fourth region, and as in Griffin et al. (2010) and Jacobs (2016), that is included beginning in July 1994.

The result is a comprehensive global dataset spanning 56,171 unique stocks and more than 7.5 million stock-month observations. Table 1 shows the descriptive statistics for the stocks in the final sample.

[Table 1 about here.]

2.2 Factor models and factors

I analyze eight different factor models comprising eleven separate factors. I adopt a common set of construction rules for the set of global factors to improve comparability. The market factor, RMRF, is the value-weight market portfolio return minus the risk-free rate (one-month Treasury bill rate, obtained from Kenneth French's website). The remaining factors are constructed based on 2 x 3 sorts on size and the respective factor criteria. With regard to the size breakpoints in the 2 x 3 sorts, I follow the common approach of Fama and French (2012, 2017) for international data: the stocks in the top 90% of the aggregate market capitalization of a country are classified as big and the stocks in the bottom 10% are classified as small. For the other sorting variable besides size, I

[3] See https://www.msci.com/market-classification for details.

calculate the breakpoints as the 30th and 70th percentiles of big stocks per country.[4]

The independent 2 x 3 sorts on size and the respective factor criteria produce six portfolios for which I compute monthly value-weight returns. The size factor (SMB, small minus big) is the average of $\text{SMB}_{\text{B/M}}$, SMB_{OP} , and SMB_{INV} which are the difference between the average returns of the three small stock and the three big stock portfolios at the intersection size and book-to-market equity, operating profitability, and investment, respectively. All other factors are the difference between the average return of the two long and the two short portfolios for the respective criteria (details of the criteria definitions are provided in Appendix A.2). The size (SMB, small minus big), the standard value (HML, high minus low), the three profitability (RMW, robust minus weak based on a certain profitability ratio), and the investment (CMA, conservative minus aggressive) factors are updated annually at the end of each June. In contrast, the momentum (WML, winners minus losers), the more timely value (HML_{m}), and the two mispricing factors, management (MGMT) and performance (PERF), are updated monthly. I use only annual accounting information, as the reporting of quarterly accounting data is not common practice in many non-U.S. countries at the beginning of my sample period and values from the fiscal year ending in the calendar year $y-1$ are only applied to predict returns from July of year y to June of year $y+1$. All returns are in U.S. dollars. Table 2 lists the applied factor models and the factors each one includes.

[Table 2 about here.]

The first model, the CAPM, uses the market factor (RMRF) as the only factor. The second model is the Fama and French (1993) three-factor model (FF3), which adds size (SMB) and value (HML) factors to the CAPM. The third model is the Fama and French (2015) five-factor model (FF5), which adds operating profitability ($\text{RMW}_{\text{OPtBE}}$) and investment (CMA) factors to the FF3 model. The fourth model is a six-factor model (FF6) that augments the FF5 with the momentum factor (WML). The fifth model is the Fama and French (2018) six-factor model (FF6_{CP}) that replaces the operating profitability factor ($\text{RMW}_{\text{OPtBE}}$) with a cash-based profitability factor ($\text{RMW}_{\text{CbOPtA}}$). The sixth model is the Hou et al. (2015) q-factor model (HXZ4), which adds size

[4] Although "Worldscope is designed for the user who needs to compare the financial information of companies from different industries and countries throughout the world" and "differences in accounting terminology, presentation, and language are minimized" by defining standardized data fields (Thomson Reuters, 2013, p. 30), differences due to country-specific accounting practices cannot be ruled out. Therefore, I calculate country-specific breakpoints instead of regional-specific breakpoints as Fama and French (2012, 2017) do. In Section 5, I also calculate factors based on breakpoints per region and find my conclusions unchanged.

(SMB), investment (CMA), and profitability (RMW_{ROE}) factors to the CAPM.[5] The seventh model is the Stambaugh and Yuan (2017) mispricing factor model (SY4), which includes market (RMRF), size (SMB), and two composite mispricing factors, management (MGMT) and performance (PERF). The final model is the six-factor model ($FF6_{CP,m}$) proposed in Barillas et al. (2019), which substitutes the HML factor in the $FF6_{CP}$ for a monthly updated value factor (HML_m).

[Table 3 about here.]

Table 3 presents summary statistics for the monthly factor returns between July 1990 and October 2018, including mean, standard deviation, and t-statistic. All factors exhibit positive and mostly sizable average returns. The factor with the highest average return is WML, followed by the two value factors, HML and HML_m. The size factor, SMB, has the smallest average return and, other than the market factor, is the only one with a t-statistic below 2. All of the other factors, aside from the profitability factor based on ROE, RMW_{ROE}, and the investment factor CMA have t-statistics greater than 3, the critical t-value recommended in Harvey et al. (2016). The cash-based operating profitability factor, RMW_{CbOPtA}, has the highest t-statistic, partly due to its low standard deviation, which is lowest together with the standard deviation of the operating profitability factor, RMW_{OPtBE}. Overall, the patterns documented for markets outside the U.S. are quite similar to the patterns reported for the U.S. in Barillas et al. (2019). This result already alleviates concerns that the documented patterns for the U.S. are due to chance.

[Table 4 about here.]

Table 4 shows the correlations among the factor returns. As expected, factor returns in the same category, such as the two value factors and the three profitability factors, tend to be highly correlated. Furthermore, returns for the two composite mispricing factors, MGMT and PERF, show also high correlations with factors based on the underlying mispricing measures: the management factor, MGMT, is highly correlated with the investment factor, and the performance factor, PERF, is highly correlated with the momentum and profitability factors. As in Asness and Frazzini (2013) and Barillas et al. (2019), and as expected by the factor construction, the momentum factor, WML,

[5] Due to my factor construction convention, the size and investment factors in the HXZ4, FF5, FF6 models are identical although Hou et al. (2015) construct their size, investment, and profitability factors from a triple 2 x 3 x 3 sort on size, ROE, and investment-to-assets. Furthermore, they use quarterly data to measure investment and profitability.

is far more negatively correlated with the timely value factor, HML_m, than with the standard value factor, HML. The strongly negative correlation between HML_m and WML already indicates potential diversification gains.

3 Methodology

I compare the various factor models in three different ways. First, I compute mean-variance efficient (maximum Sharpe ratio) portfolios for the different factor models. Second, I conduct pairwise tests of equality of squared Sharpe ratios. Finally, I present the result of factor spanning tests. This section describes the three testing procedures.

As in Hirshleifer, Hsu, and Li (2013), Ball et al. (2016), Daniel et al. (2019), or Keloharju, Linnainmaa, and Nyberg (2020), I compute and compare the maximum Sharpe ratios that can be achieved with different set of factors. Differences in these Sharpe ratios measure the economic significance of the results, i.e., they quantify how much a hypothetical investor would benefit by having certain factors in an investment opportunity set (ignoring transaction costs and short-selling constraints). Furthermore, I present the mean-variance efficient tangency portfolio weights to provide insights into the relative importance of each of the factors within a model.

However, comparing mean-variance efficient tangency portfolios as described above does not provide statistical evidence. Therefore, I conduct pairwise tests of equality of the models' squared Sharpe ratios as in Barillas et al. (2019).[6] First, I calculate the differences between the bias-adjusted sample squared Sharpe ratios for various pairs of factor models. The squared Sharpe ratio for each model is modified to be unbiased for small samples under joint normality by multiplying it by $(T - K - 2)/T$ and subtracting K/T, where T is the number of return observations and K is the number of factors. Second, I compute p-values for the test of equality of the squared Sharpe ratios.

For nested models (i.e., all of the factors in one model are contained in the other model), I test whether the squared Sharpe ratio for the model with more factors is higher than the squared Sharpe ratio of the model with fewer factors. This becomes equivalent to testing whether the factors in the larger model that are not also in the smaller model have significant alphas when regressed on the smaller model. For example, because the CAPM is nested in the FF3, I test whether the

[6] I am grateful to Cesare Robotti for the provision of the Matlab code to conduct these model comparison tests.

CAPM alphas of SMB and HML are zero. The statistical test for multiple excluded factors, as in this example, is the application of the (GRS) test in Gibbons et al. (1989). For a single excluded factor, the corresponding test is whether the intercept in the regression of the excluded factor on the nested model is significantly different from zero.

When the models are non-nested (i.e., each model contains factors not included in the other model), I conduct a sequential test. I first determine whether the squared Sharpe ratio of the model that includes all of the factors from both models is larger than for the model composed of the common factors.[7] A GRS test evaluates whether the alphas of the non-common factors based on the models with the common factors are zero. For example, because the HXZ4 and SY4 models are non-nested, I conduct the GRS test of the non-common factors, RMW_{ROE}, CMA, MGMT, and PERF, on the common factors model (RMRF + SMB). If this test cannot be rejected, then the evidence is consistent with the notion that the squared Sharpe ratio of each model equals that of the model consisting of the common factors, and that the common factors model is as good as the models that add the non-common factors. Thus, the non-common factors are redundant and the two non-nested models are equivalent under this null. If this preliminary test is rejected, some or all of the non-common factors are not redundant and contribute to an increase in the squared Sharpe ratio compared to the common-factors model. However, the preliminary test does not tell us which non-nested model has a higher squared Sharpe ratio, and equality may or may not hold for the two models. In this case, I use the direct test of squared Sharpe ratio equality as outlined in Proposition 1 in Barillas et al. (2019). The p-value in this direct test is calculated as the (bias-adjusted) squared Sharpe ratio difference divided by its standard error. The standard error of the squared Sharpe ratio difference is the square root of the asymptotic variance divided by the number of monthly observations.

Finally, I present the result of factor spanning tests to better understand why the non-common factors of the best performing model add information to the remaining factors. In a spanning test, I conduct a time-series regression of the returns of a given factor on the factor returns from a benchmark factor model that does not include that factor. If this regression shows sizable and statistically significant alphas, the left-hand side factor contains important information that is not

[7] As all investigated models include the market and size factors, the models have at least two common (overlapping) factors.

covered by the benchmark model. This will not give new insights into *which* model is the best but will provide some insights into *why* the model performs well.

4 Empirical Results

4.1 Comparing models based on maximum Sharpe ratios

To measure the economic significance of my findings, I follow Hirshleifer et al. (2013), Ball et al. (2016), Daniel et al. (2019), and Keloharju et al. (2020) and compute the ex post maximum Sharpe ratios associated with various factor combinations. Differences in these Sharpe ratios measure the economic significance of the results, i.e., they quantify how much a hypothetical investor would benefit by having certain factors in an investment opportunity set (ignoring transaction costs and short-selling constraints). Table 5 displays the tangency portfolio weights and the corresponding maximum Sharpe ratios.

[Table 5 about here.]

The (annualized) Sharpe ratio for the international market portfolio (CAPM) is 0.21, and increases to 0.83 when I extend the factor set by adding the classic size and value factors to create the Fama and French (1993) three-factor model (FF3). Adding two more factors to obtain the Fama and French (2015) five-factor model (FF5) leads to a maximum Sharpe ratio of 1.41. An investor who holds the previously mentioned factors with the appropriate weightings would benefit slightly by adding a momentum factor (FF6), as the Sharpe ratio increases to 1.58. Furthermore, replacing the operating profitability factor in the FF6 model with a cash-based profitability factor to create the FF6_{CP} model as in Fama and French (2018) would further increase the Sharpe ratio to 1.97. As an intermediate result, I conclude that the step by step extensions of the CAPM by Fama and French (1993, 2015, 2018) successively increase the Sharpe ratios in markets outside the U.S. and therefore result in more powerful factor models. This result alleviates concerns that the findings in Fama and French (1993, 2015, 2018) for the U.S. market are due to chance or to data snooping. Next, I investigate how the Fama and French models perform compared to competing models.

Switching from the Fama and French (2015, 2018) five- or six-factor models to the models proposed in Hou et al. (2015) or Stambaugh and Yuan (2017) would result in lower maximum

Sharpe ratios of 1.23 (HXZ4) and 1.43 (SY4), respectively. Therefore, using these latter models would not lead to more attractive investment opportunity sets. In contrast, replacing the value factor in the Fama and French (2018) six-factor model with a monthly updated value factor as in Barillas et al. (2019) would be highly beneficial. The resulting six-factor model ($FF6_{CP,m}$) achieves the highest maximum Sharpe ratio with a value of 2.36, which is a substantial improvement of about 20% and 50% compared to the $FF6_{CP}$ and FF6 models, respectively. Compared to the FF3, FF5, HXZ4, and SY4 models the improvement is even larger.

Figure 1 plots the efficient frontiers for the models as mentioned above, their tangency portfolios, and the underlying factors. It shows that the $FF6_{CP,m}$ model not only spans the dominating tangency portfolio but also dominates the competing models for different levels of risk, as for any given level of risk (standard deviation) no other model provides a higher return.

[Figure 1 about here.]

These results document that the six-factor model proposed in Barillas et al. (2019) spans the highest Sharpe ratio among the models investigated here. However, the analysis does not provide statistical evidence. Therefore, I conduct pairwise tests of equality of squared Sharpe ratios as in Barillas et al. (2019), with results shown in Table 6.

[Table 6 about here.]

Panel A in Table 6 shows the differences between the (bias-adjusted) sample squared Sharpe ratios (column model minus row model) for various pairs of models. Panel B, reports p-values for the tests of equality of the squared Sharpe ratios. The estimate of the squared Sharpe ratio for each model is modified to be unbiased as in Barillas et al. (2019). Similarly, the p-values are computed differently depending on whether the models to be compared are nested or non-nested.[8]

The main findings can be summarized as follows: First, the results show that the CAPM is dominated by all other models, with significance at the 1% level. Similarly, the FF3 model is outperformed by the SY4 model at the 5% level, and by all remaining models at the 1% level, except for the HXZ4 model which has a higher squared Sharpe ratio than the FF3 model, but the difference is not statistically different. The SY4 model produces a higher Sharpe ratio than both the FF5 and HXZ4 models, but the differences between the three models are not significant.

[8] See also Section 3.

When I augment the FF5 model with the momentum factor, the resulting FF6 model outperforms the HXZ4 model at the 5% level and the FF5 model at the 1% level but not the SY4 model at conventional significance levels. However, when substituting the operating profitability factor in the FF6 model for a cash-based profitability factor, the resulting FF6_{CP} outperforms all models mentioned above at least at the 10% level. The last column highlights the importance of the choice of the value factor in the Fama and French (2018) six-factor model. The $\text{FF6}_{\text{CP,m}}$ six-factor model that includes a cash-based profitability factor and a timely value factor as shown in Barillas et al. (2019) dominates all other models and the squared Sharpe ratios are different, as indicated by the associated p-values that are virtually zero.

4.2 Value is not redundant

Given the importance of the choice of the value factor as shown in the previous subsection, I conduct spanning regressions of the monthly updated value factor returns (HML_{m}) on the returns of the factors in the models that do not contain this factor. The goal of this analysis is to obtain more insights into why the monthly updated value factor is so powerful.

[Table 7 about here.]

Table 7 presents the results of a series of spanning tests where I regress the returns of the value factor that is updated monthly on the factor returns of the following models: (i) the Fama and French (2015) five-factor model augmented with momentum (FF6), (ii) the Fama and French (2018) six-factor model (FF6_{CP}) that substitutes the operating profitability factor in the FF6 model for a cash-based profitability factor, (iii) the Hou et al. (2015) q-factor model (HXZ4), (iv) the Stambaugh and Yuan (2017) mispricing model (SY4), and (v) all factors jointly. In all spanning regressions, HML_{m} exhibits sizable and highly significant alphas. The alpha ranges from 27 bps (t-value of 7.08) for the Fama and French (2015) five-factor model augmented with momentum to 58 bps (t-value of 5.77) for the Stambaugh and Yuan (2017) mispricing model. Even when the spanning regression includes all of the factors, the alpha of the monthly updated value factor is 34 bps (t-value of 8.80). Further analysis of the results of the spanning tests shows that HML_{m} exhibits a positive and highly significant alpha, even with a very large loading (0.88) on the classic HML factor that on a standalone basis has a similar return as HML_{m}. The reason is the negative loadings on the momentum, cash-based profitability, and performance factors. These three factors exhibit high and significant returns standalone, and their negative correlations with the monthly updated value factor explain why HML_m is such a strong contributor compared to HML (see also Table 4).

Figure 2 shows the cumulative performance and the spanning alpha of HML_m over time. I use the full sample estimates of the spanning betas to estimate the monthly alphas (from the last model in Table 7).

[Figure 2 about here.]

HML_m delivers a strong performance over the full sample period, but has a rather flat period in the mid-1990s and sharpe drawdown during the dot-com bubble period toward the end of the 1990s, followed by a substantial reversal in the bear market of the early 2000s. I also note that the standalone performance of the monthly updated value factor has been flat over the decade of the 2010s. However, when controlling for other factors, the spanning test alpha for the monthly updated value factor shows a reliable and robust performance over time.

5 Robustness

To check the robustness of the results in the previous section, I repeat the maximum Sharpe ratio analysis (i) across five-year subperiods and in out-of-sample tests, (ii) across global subregions (Asia, Europe, Japan, and emerging markets), and (iii) using various methodological changes.

5.1 Subperiod and out-of-sample analysis

The results presented in the previous section are based on the full sample period from July 1990 to October 2018. However, factor returns, and therefore model performance, vary over time. Fama and French (2018) also find that their model rankings vary slightly over time. To investigate the consistency of my model ranking over time, I rerun the analysis of Table 5 for non-overlapping subperiods of five years.[9] Table 8 presents the maximum ex post annualized Sharpe ratios that can be achieved using various factor combinations for the six subperiods.

[9]The only methodological difference is that I require positive factor weights for the tangency portfolio, to avoid having the signs of the factor weights switch for different subsample periods. However, removing this restriction does not alter the main conclusion.

[Table 8 about here.]

The results in Table 8 show that the $FF6_{CP,m}$ six-factor model proposed in Barillas et al. (2019) achieves the highest Sharpe ratio in five out of the six subperiods. Only for the subperiod from July 2010 to June 2015 does the Fama and French (2018) six-factor model with the classical value factor ($FF6_{CP}$) outperform, with an annualized Sharpe ratio of 4.47. The $FF6_{CP,m}$ six-factor model ranks second in this subperiod, with a slightly lower annualized Sharpe ratio of 4.35.

My conclusions thus far have been based entirely on ex post (in-sample) Sharpe ratios. As Barillas and Shanken (2018) and Kan, Wang, and Zheng (2019) highlight, there is likely to be an upward bias in the sample Sharpe ratios of the models identified as the best performers. It is therefore interesting to investigate the out-of-sample performance of the models. I apply an estimation period for determining tangency portfolio weights for the different models. These factor weights are then applied to calculate factor returns in out-of-sample performance evaluations. Table 9 presents the out-of-sample annualized Sharpe ratios for the different models for rolling estimation periods of 36, 60, 120, and 180 months.

[Table 9 about here.]

The results in Table 9 again show that the $FF6_{CP,m}$ six-factor model proposed in Barillas et al. (2019) achieves the highest Sharpe ratio for each of the four rolling estimation windows. The annualized out-of-sample Sharpe ratios are all still above 2, and are therefore close to the maximum Sharpe ratio delivered by the $FF6_{CP,m}$ model for the out-of-sample period as shown in the last column.[10]

5.2 Regional analysis

Jacobs and Müller (2018) not only document the multidimensionality of individual stock returns in global markets but also show that the most important return predictors differ across regions. To investigate the consistency of the model ranking across regions, I rerun the analysis in Table 5 for the global regions of Asia Pacific (excluding Japan), Europe, Japan, and emerging markets. Table 10 presents the annualized maximum ex post Sharpe ratios that can be achieved by various

[10] The maximum Sharpe ratios differ per row as the out-of-sample periods have different lengths depending on the length of the rolling estimation window.

factor combinations for the four subregions.

[Table 10 about here.]

The results in Table 10 are striking. The $FF6_{CP,m}$ six-factor model in Barillas et al. (2019) produces the highest Sharpe ratio for all four subregions. The difference between Sharpe ratios across the models is most pronounced for Japan while the difference is smallest for Asia Pacific, where the Fama and French (2015) five-factor model augmented with momentum (FM6) achieves a Sharpe ratio that is only 4% smaller than the one of the $FF6_{CP,m}$ model.

5.3 Methodological changes

In this final subsection, I perform a range of tests to examine the robustness of the results to plausible variations in the research design. First, I calculate region-specific breakpoints instead of country-specific breakpoints to allocate stocks into the 2 x 3 portfolio sorts. Second, I remove micro caps from my sample. Micro caps are defined as the smallest stocks that comprise 3% of the aggregated market capitalization per country. Although micro stocks represent only 3% of the total market capitalization, they account about 60% of the number of stocks. Third, I replace the ROE profitability factor with a cash-based profitability factor in the HXZ4 model. Next, I exclude financial firms (four-digit Industry Classification Benchmark codes starting with 8) from my sample. Finally, I include failure probability and O-Score in the composite score for the performance factor (PERF) of the Stambaugh and Yuan (2017) mispricing model. Table 11 presents annualized Sharpe ratios under these methodological changes compared to the base case results shown in Table 5.

[Table 11 about here.]

The results in Table 11 underline the robustness of my main conclusion, as the $FF6_{CP,m}$ model achieves the highest Sharpe ratio under all variations. Calculating region-specific instead of country-specific portfolio breakpoints produces a lower maximum Sharpe ratio for the $FF6_{CP,m}$ model. However, this Sharpe ratio is still the highest among all competing models. Removing micro caps from the sample, as suggested in Hou et al. (2018), decreases the Sharpe ratios for all models but only slightly, as all factors are market cap-weighted and breakpoints are already based on big stocks. Substituting the ROE profitability factor with a cash-based profitability factor in the HXZ4 model leads to a substantially higher maximum Sharpe ratio for this model. However, it ranks only third behind the $\mathrm{FF6_{CP}}$ and $\mathrm{FF6_{CP,m}}$. Excluding financial firms slightly increases the performance of the $\mathrm{FF6_{CP,m}}$ model, while the alternative performance factor definition only slightly decreases the performance of the Stambaugh and Yuan (2017) mispricing model.

6 Conclusion

To my knowledge, this study is the first to compare a broad set of both long-standing and recently proposed factor models using a comprehensive global sample of stock price returns. More, specifically, I investigate the following models for 50 non-U.S. markets from 1990 to 2018: (i) the CAPM, (ii) the Fama and French (1993) three-factor model, (iii) the Fama and French (2015) five-factor model, (iv) the Fama and French six-factor model (the five-factor model augmented with a momentum factor), (v) the Fama and French (2018) six-factor model that replaces the operating profitability factor with a cash-based profitability factor, (vi) the Hou et al. (2015) q-factor model that includes profitability and investment factors in addition to market and size factors, (vii) the Stambaugh and Yuan (2017) mispricing model that extends the CAPM by adding a size factor and two composite mispricing factors, management and performance, and (viii) the six-factor model proposed in Barillas et al. (2019) that substitutes the classic value factor in the Fama and French (2018) six-factor model for a monthly updated value factor. Therefore, I comprehensively extend the evidence for the U.S. market shown in Barillas et al. (2019) by comparing common factors models across markets globally.

My main findings are summarized as follows. First, I document that all employed factors exhibit positive and, with the exception of the market and size factors, significant average returns. Furthermore, the documented patterns are quite similar to the ones reported in Barillas et al. (2019) for the U.S. market. Second, I find that the six-factor model proposed in Barillas et al. (2019) produces a maximum Sharpe ratio that is substantially and significantly higher than those of competing factor models. Therefore, I conclude that a six-factor model that includes a cash-based profitability factor, a momentum factor, and a value factor that is updated monthly dominates the factor models in Fama and French (1993, 2015, 2018), Hou et al. (2015), and Stambaugh and Yuan (2017) for international markets, which is consistent with the U.S. evidence provided in Barillas et al. (2019). Third, I demonstrate that the main problem for the dominated factors models is that they cannot explain the monthly updated value factor. Hence, in contrast to Fama and French (2015), I conclude that the value factor is not redundant when it is updated monthly. Finally, my results are robust in out-of-sample tests, across five-year subperiods, across different regions (Asia, Europe, Japan, and emerging markets), and to various changes in methodology.

The documented results have important implications for defining a powerful empirical asset pricing model. Since my findings confirm the evidence for the U.S. market in Barillas et al. (2019), I conclude that their six-factor model robustly outperforms competing factor models around the globe and that this affirmative result alleviates the concerns that the findings for the U.S. could be due to chance. It follows that the six-factor model in Barillas et al. (2019) should be the preferred benchmark model for future research, both for the U.S. and in other markets. A powerful factor model should also be better able to reduce the number of anomalies in the cross-section of stock returns. For instance, Ball et al. (2016) demonstrate that a model that incorporates a cash-based profitability factor subsumes the accrual anomaly, while the Fama and French (2015) five-factor and the Hou et al. (2015) q-factor models, both of which contain profitability factors based on measures that include accruals, cannot do so (cf., Fama and French, 2016; Hou et al., 2015). Similarly, a model that contains both a monthly updated value factor and a momentum factor raises the bar for alternative value factors. An example would be a valuation multiple that is updated monthly and shows anomalous returns relative to models that include only the standard value factor and the momentum factor. However, the anomalous returns could be due to the monthly updating frequency, while the multiple itself may not carry any unique information.

This study focuses on eight global factor models. Barillas and Shanken (2017) show that when *comparing models*, the choice of test assets is irrelevant. Therefore, the only relevant comparison is how well models price the factors of competing models not included. However, Green, Hand, and Zhang (2017) find that 12 out of 94 characteristics that they investigate are reliably independent determinants of stock returns and that 11 of the 12 independent characteristics are not included in prominent benchmark models. Similarly, while exploring 161 cross-sectional predictors, Jacobs and Müller (2018) find that the most predictive variables for their global sample differ from the ones in benchmark models. Unfortunately, neither Green et al. (2017) nor Jacobs and Müller (2018) include the six-factor model proposed in Barillas et al. (2019) within their set of benchmark models. Examining how the Barillas et al. (2019) six-factor model can explain the return of these *outside* *benchmark model factors* would be an interesting extension of this research. I leave this topic for future research.

A Appendix

A.1 Datastream sample definition

Constituent lists

Datastream comprises three types of constituent lists: (1) research lists, (2) Worldscope lists, and (3) dead lists. By using dead lists, I ensure to obviate any survivorship bias. For every country, I use the union of all available lists and eliminate any duplicates. As a result, I have one remaining list for every country, which can subsequently be used in the static filter process. Table A.1 and Table A.2 provide an overview of the constituent lists for developed markets and emerging markets, respectively, used in my study.

[Table A.1 about here.]

[Table A.2 about here.]

Static screens

I restrict my sample to common equity stocks by applying several static screens as shown in Table A.3. Screen (1) to (7) are straightforward to apply and common in the literature.

[Table A.3 about here.]

Screen (8) is related to, among others, the following work: Ince and Porter (2006), Campbell, Cowan, and Salotti (2010), Griffin et al. (2010), Karolyi, Lee, and van Dijk (2012). The authors provide generic filter rules in order to exclude non-common equity securities from Thomson Reuters Datastream. I apply the identified keywords and match them with the security names provided by Datastream. A security is excluded from the sample in case a keyword coincides with part of the security name. The following three Datastream items store security names and are applied for the keyword filters: “NAME”, “ENAME”, and “ECNAME”. Table A.4 gives an overview of the keywords used.

[Table A.4 about here.]

In addition, Griffin et al. (2010) introduce specific keywords for individual countries. Thus, the keywords are applied to the security names of single countries only. Exemplary, German security names are parsed to contain the word "GENUSSSCHEINE", which declares the security to be non-common equity. In Table A.5, I give an overview of country-specific keyword deletions conducted in my study.

[Table A.5 about here.]

Dynamic screens

For the securities, remaining from the static screens above, I obtain return and market capitalization data from Datastream and accounting data from Worldscope. Several dynamic screens that are common in the literature were installed in order to account for data errors mainly within return characteristics. The dynamic screens are shown in Table A.6.

[Table A.6 about here.]

A.2 Factor construction and definition

I follow Fama and French (1993, 2012, 2015, 2017) in constructing my set of factors. The market factor, RMRF, is the value-weight market portfolio return minus the risk-free rate (one-month Treasury bill rate, obtained from Kenneth French's website).

The remaining factors are constructed based on 2 x 3 sorts on size and the respective factor criteria. With regard to the size breakpoints in the 2 x 3 sorts, I follow the common approach of Fama and French (2012, 2017) for international data: the stocks in the top 90% of the aggregate market capitalization of a country are classified as big and the stocks in the bottom 10% are classified as small. For the other sorting variable besides size, I calculate the breakpoints as the 30^{th} and 70^{th} percentiles of big stocks per country.[11]

The independent 2 x 3 sorts on size and the respective factor criteria produce six portfolios for which I compute monthly value-weight returns. The size factor (SMB, small minus big) is the

[11] Although "Worldscope is designed for the user who needs to compare the financial information of companies from different industries and countries throughout the world" and "differences in accounting terminology, presentation, and language are minimized" by defining standardized data fields (Thomson Reuters, 2013, p. 30), differences due to country-specific accounting practices cannot be ruled out. Therefore, I calculate country-specific breakpoints instead of regional-specific breakpoints as Fama and French (2012, 2017) do. In Section 5, I also calculate factors based on breakpoints per region and find my conclusions unchanged.

average of $\text{SMB}_{\text{B/M}}$, SMB_{OP} , and SMB_{INV} which are the difference between the average returns of the three small stock and the three big stock portfolios at the intersection size and book-to-market equity, operating profitability, and investment, respectively. All other factors are the difference between the average return of the two long and the two short portfolios for the respective criteria. Depending on the factor, I update the portfolios monthly at the end of month t or yearly at the end of June of year y and the factor portfolios are long stocks with high or low values in the respective factor criteria and vice versa for the short leg (details see below). Following the original studies, I exclude financial firms when calculating cash-based operating profits-to-assets, gross profits-to-assets, investment-to-assets, net operating assets, and operating accruals. Firms with negative book values are excluded when sorting on both book-to-market equity ratios, operating profits-to-book equity, and return on equity. All returns are in U.S. dollars.

For each variable, I describe the detailed variable definition and used Worldscope items below. Values from the balance sheet, the income statement, and the statement of cash flows from the fiscal year ending in the calendar year $y-1$ are only applied to predict returns from July of year y to June of year $y+1$.

Size factor: SMB

The size factor (SMB, small minus big) is based on the market capitalization of a stock. Market capitalization (Datastream item MV) is price times the number of shares outstanding.

Value factors: HML and HML_m

Both value factors (HML and HML_m, high minus low) are based on book-to-market equity ratios. The nominator of both book-to-market ratios is book equity. Book equity is defined as common equity (Worldscope item WC03501) plus deferred taxes (WC03263, zero if missing). Market capitalization (Datastream item MV) is the denominator. HML uses market capitalization as measured at the end of December of year $y-1$ to predict returns from July of year y to June of year $y+1$. HML_m uses the most recent market capitalization (cf. Asness and Frazzini, 2013) and is updated monthly.

Profitability factors: $\mathbf{RMW_{OPtBE}}$, $\mathbf{RMW_{ROE}}$, and $\mathbf{RMW_{CbOPtA}}$

The profitability factors (RMW, robust minus weak) are based on return on equity (ROE), operating profits-to-book equity (OPtBE), and cash-based operating profits-to-asset (CbOPtA), respectively. I measure return on equity as earnings before extraordinary items (WC01551) divided by book equity. I measure operating profits-to-book equity as operating income (WC01250) divided by book equity. As in Ball et al. (2016), cash-based operating profits-to-asset is operating profits converted to a cash basis divided by total assets (WC02999). Following Ball, Gerakos, Linnainmaa, and Nikolaev (2015), operating profits is net sales or revenues (WC01001) minus cost of goods sold (WC01501) minus selling, general, and administrative expenses (WC01101), excluding research and development expense (WC01201). The cash-based adjustment is the year-on-year change in deferred income (WC03262), plus change in accounts payable (WC03040), plus change in accrued expenses (WC03054 + WC03069), minus change in accounts receivable (WC02051), minus change in inventory (WC02101), minus prepaid expenses (WC02140), all divided by total assets. All changes are set to zero if missing. Book equity is defined as common equity (WC03501) plus deferred taxes (WC03263, zero if missing). The profitability factors are long firms with high profitability and updated yearly.

Investment factor: CMA

The investment factor (CMA, conservative minus aggressive) is based on asset growth. As in Cooper, Gulen, and Schill (2008), I measure asset growth in June of year y as the percentage change in total assets (WC02999) from the fiscal year ending in the calendar year $y-2$ to the fiscal year ending in the calendar year $y-1$. The investment factor is long firms with low asset growth and updated yearly.

Momentum factor: WML

The momentum factor (WML, winner minus loser) is based on the cumulated total stock return (calculated via the total return index, Datastream item RI) from month $t-12$ to month $t-2$, where t is the month of the forecasted return. Skipping the last month is standard in the momentum literature to avoid an overlap with the short-term reversal effect as documented by Jegadeesh (1990).

The momentum factor is long firms with high past returns and updated monthly.

Mispricing factors: MGMT and PERF

The management (MGMT) and performance (PERF) factors are two composite mispricing factors and are constructed following Stambaugh and Yuan (2017). Furthermore, I follow the necessary adjustments for my international setting as discussed in Jacobs (2016) and Lu, Stambaugh, and Yuan (2017). For each underlying mispricing measure, and for each measure-month-country combination, I first rank stocks and compute stock percentile values in a way that the presumably most underpriced stock, i.e., highest average abnormal return, as reported in the literature, receives the lowest value and vice versa. The composite score for MGMT is computed as the average percentile of asset growth, composite equity issues, investment-to-assets, net stock issues, net operating assets, and operating accruals. The composite score for PERF is computed as the average percentile of gross profitability, six-month momentum, and return on assets. In contrast to Stambaugh and Yuan (2017) and in line with Lu et al. (2017), I do not include failure probability and O-Score in my main analysis as these measures are calculated based on models parameterized for the U.S. market. In Section 5, I include these two measures in the calculation of the performance factor and find my conclusions unchanged. For both MGMT and PERF, I require at least three individual anomalies to be non-missing to create the composite mispricing score. The two mispricing composites are long firms with low mispricing scores and updated monthly.

Six-month momentum

Six-month momentum is calculated as the cumulated total stock return (calculated via the total return index, Datastream item RI) from month $t-6$ to month $t-2$, where t is the month of the forecasted return. Skipping the last month is standard in the momentum literature to avoid an overlap with the short-term reversal effect as documented by Jegadeesh (1990).

ROA: Return on assets

Return on assets is defined as earnings before extraordinary items (WC01551) divided by total assets (WC02999).

GP/A: Gross profits-to-assets

As in Novy-Marx (2013), gross profits-to-assets is net sales or revenues (WC01001) minus cost of goods sold (WC01501), both divided by total assets (WC02999).

OA: Operating accruals

Following Sloan (1996), I define operating accruals as the change in operating working capital minus depreciation, depletion, and amortization (WC01151, zero if missing); all deflated by total assets (WC02999). Change in operating working capital is the change in current assets (WC02201) minus change in cash and short-term investments (WC02001), minus change in current liabilities (WC03101), plus change in debt in current liabilities (WC03051, zero if missing), plus change in income taxes payable (WC03063, zero if missing). Operating accruals in June of year y are measured from the fiscal year ending in the calendar year $y-2$ to the fiscal year ending in the calendar year $y-1$.

NOA: Net operating assets

As in Hirshleifer, Hou, Teoh, and Zhang (2004), net operating assets in June of year y are defined as operating assets minus operating liabilities for the fiscal year ending in the calendar year $y-1$; all deflated by total assets (WC02999) for the fiscal year ending in the calendar year $y-2$. Operating assets is total assets (WC02999) minus cash and short-term investment (WC02001). Operating liabilities is total assets minus short-term and long-term debt (WC03255), minus minority interest (WC03426, zero if missing), minus preferred stock and common equity (WC03995).

NSI: Net stock issues

Following Pontiff and Woodgate (2008), I measure net stock issues as the difference in the natural logs of split-adjusted shares outstanding in month $t-1$ and month $t-13$. Split-adjusted shares outstanding are calculated as shares outstanding (Datastream item NOSH) divided by the adjustment factor (Datastream item AF). I apply a 12-month lookback window and update the variable monthly.

CEI: Composite equity issuance

Similar to Daniel and Titman (2006), composite equity issuance is defined as the growth rate in the market capitalization not attributable to the total stock return R: $log(MC_{t-1}/MC_{t-13}) - R_{(t-13,t-1)}$. For the portfolio formation at the end of month $t-1$, $R_{(t-13,t-1)}$ is the cumulative log return (calculated via the total return index, Datastream item RI) from month $t-13$ to month t and MC_{t-1} is the market capitalization (Datastream item MV) from the end of month t. Equity issuance such as stock issues, stock option plans, and share-based compensations and acquisitions increase the composite equity issuance, whereas share repurchases or cash dividends reduce the composite equity issuance. I apply a 12-month lookback window and update the variable monthly.

I/A: Investment-to-assets

As in Lyandres, Sun, and Zhang (2008), I measure investment-to-assets in June of year y as the change in gross property, plant, and equipment (WC02301) plus the annual change in inventories (WC02101) (both from fiscal year ending in calendar year $y-2$ to fiscal year ending in calendar year $y-1$) all divided by total assets (WC02999) of year $y-2$.

Table 1: Descriptive statistics

The table presents summary statistics for the 50 countries of my global sample. Column 2 states the market affiliation according to MSCI, with DM as Developed Markets and EM as Emerging Markets. Columns 3 and 4 report the sample start and end date for each country. Columns 5, 6 and 7 display the total, minimum, and maximum number of stocks per country, respectively. Columns 8 reports the total stock months per country. Columns 9 states the average mean size per country-month. Column 10 shows the average total size per country-month and the last column reports these values in percentage of the respective total across countries. Size is measured as market capitalization in million USD.

Country	Market	Start date	End date	Total no. stocks	Min no. stocks	Max no. stocks	Stock months	Average mean size	Average total size	Average total size in %
Australia	DM	1990-07-31	2018-10-31	3223	446	1666	388034	531	695103	3.24
Austria	DM	1990-07-31	2018-10-31	203	55	102	27058	972	72765	0.34
Belgium	DM	1990-07-31	2018-10-31	355	125	176	51136	1419	209544	0.98
Canada	DM	1990-07-31	2018-10-31	6719	1846	2733	797510	352	851310	3.96
Denmark	DM	1990-07-31	2018-10-31	377	130	210	59488	905	143990	0.67
Finland	DM	1990-07-31	2018-10-31	262	42	150	38370	1368	172446	0.80
France	DM	1990-07-31	2018-10-31	2178	637	927	262285	1725	1319828	6.14
Germany	DM	1990-07-31	2018-10-31	1696	482	1092	261791	1404	1105166	5.14
Greece	DM/EM	1994-07-29	2018-10-31	405	121	317	66946	316	76887	0.36
Hong Kong	DM	1990-07-31	2018-10-31	1801	121	1600	232840	453	422058	1.96
Ireland	DM	1993-07-30	2016-10-31	87	30	54	10881	1626	59988	0.28
Israel	DM/EM	1995-07-31	2018-10-31	823	362	564	129424	230	102754	0.48
Italy	DM	1990-07-31	2018-10-31	673	181	316	82249	1839	455907	2.12
Japan	DM	1990-07-31	2018-10-31	5449	2065	3926	1117987	1096	3589267	16.71
Netherlands	DM	1990-07-31	2018-10-31	306	77	191	42951	2687	297520	1.38
New Zealand	DM	1990-07-31	2018-10-31	314	64	134	36029	307	34399	0.16
Norway	DM	1990-07-31	2018-10-31	556	79	206	51492	842	138965	0.65
Portugal	DM/EM	1994-07-29	2018-10-31	173	43	108	18819	1051	59224	0.28
Singapore	DM	1990-07-31	2018-10-31	867	130	580	137008	518	224706	1.05
Spain	DM	1990-07-31	2018-10-31	377	124	170	48254	3340	468585	2.18
Sweden	DM	1990-07-31	2018-10-31	1387	130	766	120051	784	303756	1.41
Switzerland	DM	1990-07-31	2018-10-31	427	183	255	76434	3591	831315	3.87
U.K.	DM	1990-07-31	2018-10-31	4477	1121	1822	482175	1578	2175044	10.12

[Continued on next page]

Country	Market	Start date	End date	Total no. stocks	Min no. stocks	Max no. stocks	Stock months	Average mean size	Average total size	Average total size in %
Argentina	EM	1994-07-29	2009-06-30	118	51	81	11994	485	32333	0.15
Brazil	EM	1994-07-29	2018-10-31	293	41	199	31612	2263	299481	1.39
Chile	EM	1994-07-29	2018-10-31	251	101	154	36066	1195	142047	0.66
China	EM	1997-07-31	2018-10-31	3579	600	3428	435024	1236	2640580	12.29
Colombia	EM	1995-07-31	2017-09-29	81	30	41	6760	2592	90059	0.42
Czech Republic	EM	1997-07-31	2003-01-31	150	31	148	6725	135	12146	0.06
Egypt	EM	2001-07-31	2018-10-31	205	88	161	26736	368	47815	0.22
Hungary	EM	1997-07-31	2018-10-31	92	30	54	10300	573	21971	0.10
India	EM	1994-07-29	2018-10-31	5262	1055	3693	728058	261	767839	3.57
Indonesia	EM	1994-07-29	2018-10-31	681	144	522	89618	527	187883	0.87
Jordan	EM	2006-07-31	2009-06-30	216	168	210	6790	184	34629	0.16
Malaysia	EM	1994-07-29	2018-10-31	1294	400	987	232079	331	262595	1.22
Mexico	EM	1994-07-29	2018-10-31	231	75	108	26958	1928	182015	0.85
Morocco	EM	2001-07-31	2014-06-30	88	43	76	9451	665	43255	0.20
Pakistan	EM	1994-07-29	2018-10-31	394	148	300	45071	96	24360	0.11
Peru	EM	1994-07-29	2018-10-31	156	30	69	14067	586	28460	0.13
Philippines	EM	1994-07-29	2018-10-31	308	96	242	53866	484	99997	0.47
Poland	EM	1995-07-31	2018-10-31	1083	37	760	104928	246	96595	0.45
Qatar	EM	2014-07-31	2018-10-31	44	42	44	2216	3648	155238	0.72
Russia	EM	1998-07-31	2018-10-31	555	30	257	36466	2579	451029	2.10
South Africa	EM	1995-07-31	2018-10-31	901	241	524	91539	1023	295286	1.37
South Korea	EM	1994-07-29	2018-10-31	3099	646	2134	420552	373	598373	2.79
Sri Lanka	EM	1994-07-29	2001-06-29	194	127	154	11850	11	1573	0.01
Taiwan	EM	1997-07-31	2018-10-31	2261	469	1810	322061	493	623377	2.90
Thailand	EM	1994-07-29	2018-10-31	891	311	656	131610	379	190642	0.89
Turkey	EM	1994-07-29	2018-10-31	506	152	392	85803	445	141484	0.66
United Arab Emirates	EM	2014-07-31	2018-10-31	103	78	88	4360	2392	200499	0.93
All	DM/EM	1990-07-31	2018-10-31	56171	8767	31302	7521772	818	19550186	100.00

Table 2: Overview factor models and included factors

The table presents an overview of the following factor models: The Capital Asset Pricing Model (CAPM), the Fama and French (1993) three-factor model (FF3), the Fama and French (2015) five-factor model (FF5), the Fama and French (2015) five-factor model augmented with momentum (FF6), the Fama and French (2018) six-factor model ($FF6_{CP}$) that substitutes the operating profitability factor for a cash-based profitability factor in the FF6, the Hou et al. (2015) q-factor model (HXZ4), the Stambaugh and Yuan (2017) mispricing model (SY4), as well as the six-factor model proposed in Barillas et al. (2019) ($FF6_{CP,m}$) which substitutes the classic value factor for a monthly updated value factor in the $FF6_{CP}$.

Model	included factors
CAPM	RMRF
FF3	RMRF, SMB, HML
FF5	RMRF, SMB, HML, RMW_{OPtBE}, CMA
FF6	RMRF, SMB, HML, RMW_{OPtBE}, CMA, WML
$FF6_{CP}$	RMRF, SMB, HML, RMW_{CbOPtA}, CMA, WML
HXZ4	RMRF, SMB, RMW_{ROE}, CMA
SY4	RMRF, SMB, MGMT, PERF
$FF6_{CP,m}$	RMRF, SMB, HML_m, RMW_{CbOPtA}, CMA, WML

Table 3: Summary statistics for monthly factor returns

The table presents the sample summary statistics for the following factors: SMB (small minus big) and HML (high minus low) are the size and value factor of the FF3 model. $\mathrm{HML_m}$ is the monthly updated value factor as in Asness and Frazzini (2013). $\mathrm{RMW_{OPtBE}}$ (robust minus weak based on operating profitability) and CMA (conservative minus aggressive) are the profitability and investment factors in the FF5 model. $\mathrm{RMW_{ROE}}$ and $\mathrm{RMW_{CbOPtA}}$ are the profitability factors based on ROE and cash-based operating profitability as in the HXZ4 and $\mathrm{FF6_{CP}}$ models, respectively. WML (winners minus losers) is the momentum factor. Finally, MGMT (management) and PERF (performance) are the two mispricing factors in the SY4 model. The analysis is performed from 07/1990 to 10/2018.

	mean	std. dev.	t-stat.
RMRF	0.29	4.83	1.11
SMB	0.10	2.00	0.89
HML	0.44	2.02	4.00
$\mathrm{HML_m}$	0.45	2.37	3.50
$\mathrm{RMW_{OPtBE}}$	0.20	1.17	3.21
$\mathrm{RMW_{ROE}}$	0.17	1.23	2.52
$\mathrm{RMW_{CbOPtA}}$	0.31	1.17	4.86
CMA	0.23	1.48	2.84
WML	0.61	3.06	3.68
MGMT	0.28	1.57	3.32
PERF	0.43	2.20	3.61

Table 4: Correlation for monthly factor returns

The table presents the correlation matrix for our set of factors. The analysis is performed from 07/1990 to 10/2018.

	RMRF	SMB	HML	HML_m	RMW_{OPtBE}	RMW_{ROE}	RMW_{CbOPtA}	CMA	WML	MGMT
SMB	-0.03									
HML	-0.10	-0.05								
HML_m	0.05	-0.05	0.85							
RMW_{OPtBE}	-0.29	-0.18	-0.22	-0.31						
RMW_{ROE}	-0.32	-0.16	-0.25	-0.33	0.90					
RMW_{CbOPtA}	-0.38	-0.34	-0.24	-0.42	0.52	0.54				
CMA	-0.26	0.14	0.52	0.42	-0.31	-0.38	-0.27			
WML	-0.27	-0.02	-0.16	-0.59	0.31	0.29	0.40	-0.01		
MGMT	-0.21	-0.08	0.54	0.49	-0.13	-0.16	-0.14	0.71	-0.09	
PERF	-0.42	-0.16	-0.23	-0.55	0.59	0.64	0.71	-0.25	0.74	-0.20

Table 5: Maximum ex post Sharpe ratios

The table presents the annualized maximum ex post Sharpe ratios that can be achieved by various factor combinations and the individual factor weights. I analyze the following asset pricing models: The Capital Asset Pricing Model (CAPM), the Fama and French (1993) three-factor model (FF3), the Fama and French (2015) five-factor model (FF5), the Fama and French (2015) five-factor model augmented with momentum (FF6), the Fama and French (2018) six-factor model ($FF6_{CP}$) that substitutes the operating profitability factor for a cash-based profitability factor in the FF6, the Hou et al. (2015) q-factor model (HXZ4), the Stambaugh and Yuan (2017) mispricing model (SY4), as well as the six-factor model proposed in Barillas et al. (2019) ($FF6_{CP,m}$) which substitutes the classic value factor for a monthly updated value factor in the $FF6_{CP}$. The analysis is performed from 07/1990 until 10/2018.

	RMRF	SMB	HML	HML_m	RMW_{OP}	RMW_{ROE}	RMW_{CbOPtA}	CMA	WML	MGMT	PERF	SR
CAPM	1.00											0.21
FF3	0.11	0.19	0.70									0.83
FF5	0.08	0.08	0.16		0.48			0.21				1.41
FF6	0.08	0.08	0.18		0.38			0.17	0.10			1.58
$FF6_{CP}$	0.08	0.12	0.15				0.47	0.14	0.04			1.97
HXZ4	0.09	0.05				0.48		0.38				1.23
SY4	0.12	0.14								0.39	0.35	1.43
$FF6_{CP,m}$	0.07	0.12		0.21			0.43	0.07	0.11			2.36

Table 6: Tests of equality of squared Sharpe ratios

The table present pairwise tests of equality of the model's squared Sharpe ratios. I analyze the following asset pricing models: The Capital Asset Pricing Model (CAPM), the Fama and French (1993) three-factor model (FF3), the Fama and French (2015) five-factor model (FF5), the Fama and French (2015) five-factor model augmented with momentum (FF6), the Fama and French (2018) six-factor model ($FF6_{CP}$) that substitutes the operating profitability factor for a cash-based profitability factor in the FF6, the Hou et al. (2015) q-factor model (HXZ4), the Stambaugh and Yuan (2017) mispricing model (SY4), as well as the six-factor model proposed in Barillas et al. (2019) ($FF6_{CP,m}$) which substitutes the classic value factor for a monthly updated value factor in the $FF6_{CP}$. Panel A reports the difference between the (bias-adjusted) sample squared Sharpe ratios and Panel B reports the associated p-values. The analysis is performed from 07/1990 to 10/2018.

	FF3	HXZ4	FF5	SY4	FF6	$FF6_{CP}$	$FF6_{CP,m}$
Panel A: Differences in Sample Squared Sharpe Ratios							
CAPM	0.047	0.111	0.148	0.156	0.186	0.297	0.438
FF3		0.064	0.101	0.109	0.139	0.250	0.390
HXZ4			0.037	0.045	0.075	0.186	0.327
FF5				0.008	0.038	0.149	0.290
SY4					0.030	0.141	0.282
FF6						0.111	0.251
$FF6_{CP}$							0.140
Panel B: p-Values							
CAPM	0.000	0.000	0.000	0.000	0.000	0.000	0.000
FF3		0.123	0.000	0.043	0.000	0.000	0.000
HXZ4			0.197	0.319	0.040	0.006	0.000
FF5				0.877	0.001	0.023	0.000
SY4					0.483	0.015	0.000
FF6						0.051	0.001
$FF6_{CP}$							0.000

Table 7: Factor spanning tests

The table presents the results from time-series factor spanning regressions. The dependent variable is the monthly factor return of the monthly updated value factor (HML_m). The independent variables are the factors of the Fama and French (2015) five-factor model augmented with momentum (FF6), the Fama and French (2018) six-factor model ($FF6_{CP}$) that substitutes the operating profitability factor for a cash-based profitability factor in the FF6, the Hou et al. (2015) q-factor model (HXZ4), and the Stambaugh and Yuan (2017) mispricing model (SY4). I conduct a spanning regression of the monthly updated value factor (HML_m) on each of the four models mentioned above separately (first eight rows) and on all factors jointly (last two rows). The analysis is performed from 07/1990 to 10/2018.

	(Intercept)	RMRF	SMB	HML	RMW_{OPtBE}	CbOPtA	CMA	WML	RMW_{ROE}	MGMT	PERF
FF6	0.27	0.00	-0.03	0.90	0.00		0.03	-0.36			
	(7.08)	(0.04)	(-1.74)	(42.33)	(0.03)		(0.92)	(-28.63)			
$FF6_{CP}$	0.35	-0.02	-0.07	0.88		-0.21	-0.02	-0.34			
	(9.01)	(-2.34)	(-3.72)	(43.16)		(-5.22)	(-0.58)	(-27.17)			
HXZ4	0.37	0.04	-0.16				0.63		-0.34		
	(3.05)	(1.62)	(-2.68)				(6.77)		(-3.00)		
SY4	0.58	-0.06	-0.14							0.53	-0.59
	(5.77)	(-2.40)	(-2.80)							(8.15)	(-11.62)
ALL	0.34	-0.02	-0.06	0.88	0.04	-0.15	-0.10	-0.29	0.05	0.09	-0.13
	(8.80)	(-2.72)	(-3.08)	(42.54)	(0.57)	(-3.40)	(-2.45)	(-15.75)	(0.73)	(2.74)	(-3.53)

Table 8: Maximum ex post Sharpe ratios - Subperiods

The table presents the annualized maximum ex post Sharpe ratios that can be achieved by various factor combinations for six subperiods. I analyze the following asset pricing models: The Capital Asset Pricing Model (CAPM), the Fama and French (1993) three-factor model (FF3), the Fama and French (2015) five-factor model (FF5), the Fama and French (2015) five-factor model augmented with momentum (FF6), the Fama and French (2018) six-factor model ($FF6_{CP}$) that substitutes the operating profitability factor for a cash-based profitability factor in the FF6, the Hou et al. (2015) q-factor model (HXZ4), the Stambaugh and Yuan (2017) mispricing model (SY4), as well as the six-factor model proposed in Barillas et al. (2019) ($FF6_{CP,m}$) which substitutes the classic value factor for a monthly updated value factor in the $FF6_{CP}$. The subperiods comprise the full sample period from 07/1990 until 10/2018.

	07/90-06/95	07/95-06/00	07/00-06/05	07/05-06/10	07/10-06/15	07/15-10/18
CAPM	-0.03	0.32	0.05	0.25	0.73	-0.09
FF3	0.95	0.32	2.90	1.19	1.18	0.87
FF5	2.14	0.84	2.97	2.02	3.67	2.37
FF6	2.15	1.20	2.97	2.03	3.74	2.37
$FF6_{CP}$	1.62	2.02	3.35	2.33	4.37	2.37
HXZ4	1.34	0.73	2.28	1.73	3.39	1.80
SY4	0.72	1.59	1.92	1.27	3.30	1.28
$FF6_{CP,m}$	2.82	2.41	4.22	2.66	4.35	2.81

Table 9: Out-of-sample Sharpe ratios

The table presents the annualized out-of-sample Sharpe ratios for the different investigated models for rolling estimation periods of 36, 60, 120, and 180 months. I analyze the following asset pricing models: The Capital Asset Pricing Model (CAPM), the Fama and French (1993) three-factor model (FF3), the Fama and French (2015) five-factor model (FF5), the Fama and French (2015) five-factor model augmented with momentum (FF6), the Fama and French (2018) six-factor model ($FF6_{CP}$) that substitutes the operating profitability factor for a cash-based profitability factor in the FF6, the Hou et al. (2015) q-factor model (HXZ4), the Stambaugh and Yuan (2017) mispricing model (SY4), as well as the six-factor model proposed in Barillas et al. (2019) ($FF6_{CP,m}$) which substitutes the classic value factor for a monthly updated value factor in the $FF6_{CP}$. The last column ($MaxSR_{FF6_{CP,m}}$) shows the maximum (in-sample) Sharpe ratio for the $FF6_{CP,m}$ model for the respective out-of-sample period. The out-of-sample analysis is performed from 07/1993 (1995, 2000, 2005) until 10/2018.

	CAPM	FF3	FF5	FF6	$FF6_{CP}$	HXZ4	SY4	$FF6_{CP,m}$	$MaxSR_{FF6_{CP,m}}$
36	0.27	0.66	1.58	1.72	1.97	1.06	1.15	2.38	2.64
60	0.27	0.66	1.49	1.54	1.92	1.30	1.38	2.19	2.70
120	0.26	1.25	1.50	1.41	2.04	1.59	1.48	2.49	2.91
180	0.33	0.77	1.44	1.51	2.30	1.57	1.64	2.63	2.92

Table 10: Maximum ex post Sharpe ratios - regional analysis
The table presents the annualized maximum ex post Sharpe ratios that can be achieved by various factor combinations for four regions. The regions besides emerging markets are defined as in Fama and French (2012) or Fama and French (2015). I analyze the following asset pricing models: The Capital Asset Pricing Model (CAPM), the Fama and French (1993) three-factor model (FF3), the Fama and French (2015) five-factor model (FF5), the Fama and French (2015) five-factor model augmented with momentum (FF6), the Fama and French (2018) six-factor model ($FF6_{CP}$) that substitutes the operating profitability factor for a cash-based profitability factor in the FF6, the Hou et al. (2015) q-factor model (HXZ4), the Stambaugh and Yuan (2017) mispricing model (SY4), as well as the six-factor model proposed in Barillas et al. (2019) ($FF6_{CP,m}$) which substitutes the classic value factor for a monthly updated value factor in the $FF6_{CP}$. The analysis is performed from 07/1990 (07/1994) until 10/2018 for developed (emerging) markets.

	Asia Pacific	Europe	Japan	Emerging Markets
CAPM	0.37	0.34	0.04	0.17
FF3	0.83	0.56	0.60	1.17
FF5	1.24	1.33	0.66	1.43
FF6	1.73	1.48	0.68	1.67
$FF6_{CP}$	1.70	1.73	0.74	1.71
HXZ4	1.07	1.25	0.27	0.85
SY4	1.64	1.36	0.28	1.35
$FF6_{CP,m}$	1.81	2.02	1.19	2.28

Table 11: Maximum ex post Sharpe ratios - methodological changes
The table presents the annualized maximum ex post Sharpe ratios that can be achieved in plausible variations in the research design compared to the base case shown in the first column. First, I calculate region-specific breakpoints instead of country-specific breakpoints to allocate stocks into factor portfolios. Second, I remove micro caps (smallest stocks that comprise for 3% of aggregated market capitalization per country) from my sample. Third, I substitute the ROE profitability factor for a cash-based profitability factor in the HXZ4 model. Fourth, I exclude financial firms (four-digit ICB codes starting with 8) from my sample. Finally, I also include failure probability and O-Score in the composite score for the performance factor (PERF) of the Stambaugh and Yuan (2017) mispricing model. I analyze the following asset pricing models: The Capital Asset Pricing Model (CAPM), the Fama and French (1993) three-factor model (FF3), the Fama and French (2015) five-factor model (FF5), the Fama and French (2015) five-factor model augmented with momentum (FF6), the Fama and French (2018) six-factor model ($FF6_{CP}$) that substitutes the operating profitability factor for a cash-based profitability factor in the FF6, the Hou et al. (2015) q-factor model (HXZ4), the Stambaugh and Yuan (2017) mispricing model (SY4), as well as the six-factor model proposed in Barillas et al. (2019) ($FF6_{CP,m}$) which substitutes the classic value factor for a monthly updated value factor in the $FF6_{CP}$. The analysis is performed from 07/1990 until 10/2018.

	base case	regional breakpoints	removing micro caps	HXZ4 with RMW_{CbOPtA}	no financials	alt PERF definition
CAPM	0.21	0.21	0.20	0.21	0.22	0.21
FF3	0.83	0.86	0.75	0.83	0.85	0.83
FF5	1.41	1.35	1.34	1.41	1.19	1.41
FF6	1.58	1.57	1.49	1.58	1.42	1.58
$FF6_{CP}$	1.97	1.69	1.90	1.97	1.96	1.97
HXZ4	1.23	1.23	1.20	1.73	1.01	1.23
SY4	1.43	1.50	1.31	1.43	1.56	1.42
$FF6_{CP,m}$	2.36	1.94	2.30	2.36	2.39	2.36

Figure 1: Efficient Frontiers

The figure plots the efficient frontiers, the tangency portfolios (colored dots), and the underlying factors (grey dots) for the following models: The Fama and French (1993) three-factor model (FF3), the Fama and French (2015) five-factor model (FF5), the Fama and French (2015) five-factor model augmented with momentum (FF6), the Fama and French (2018) six-factor model ($FF6_{CP}$) that substitutes the operating profitability factor for a cash-based profitability factor in the FF6, the Hou et al. (2015) q-factor model (HXZ4), the Stambaugh and Yuan (2017) mispricing model (SY4), as well as the six-factor model proposed in Barillas et al. (2019) ($FF6_{CP,m}$) which substitutes the classic value factor for a monthly updated value factor in the $FF6_{CP}$. The analysis is performed from 07/1990 until 10/2018.

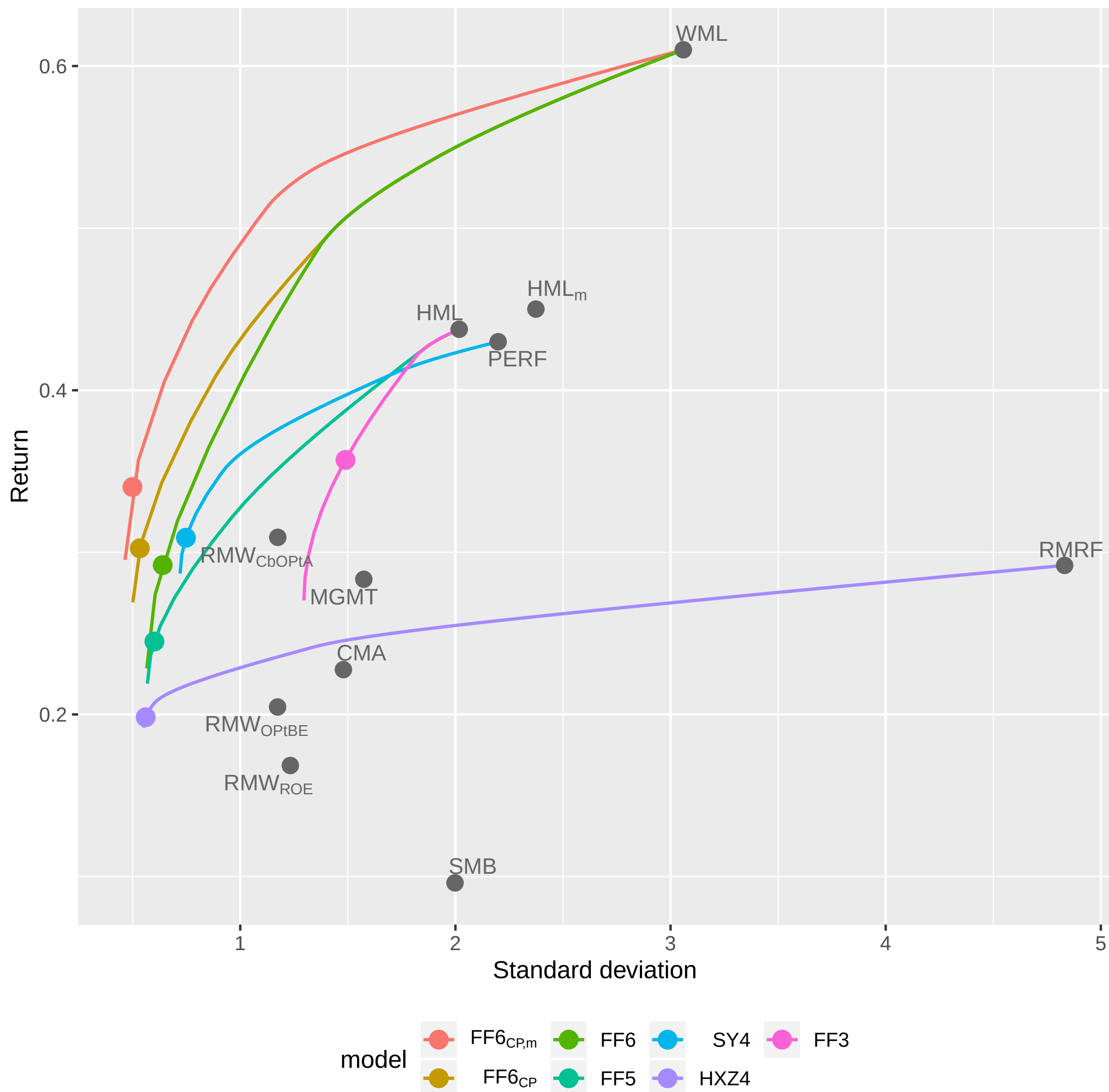

Figure 2: Cumulative performance of the monthly updated value factor
The figure plots the cumulated performance of the monthly time-series of HML_m and the spanning alpha of HML_m. The spanning alpha (last two rows in Table 7) is derived from the spanning regression of the monthly updated value factor(HML_m) on all remaining factors of the Fama and French (2015) five-factor model augmented with momentum (FF6), the Fama and French (2018) six-factor model ($FF6_{CP}$) that substitutes the operating profitability factor for a cash-based profitability factor in the FF6, the Hou et al. (2015) q-factor model (HXZ4), and the Stambaugh and Yuan (2017) mispricing model (SY4). The sample period starts in 7/1990 and ends in 10/2018.

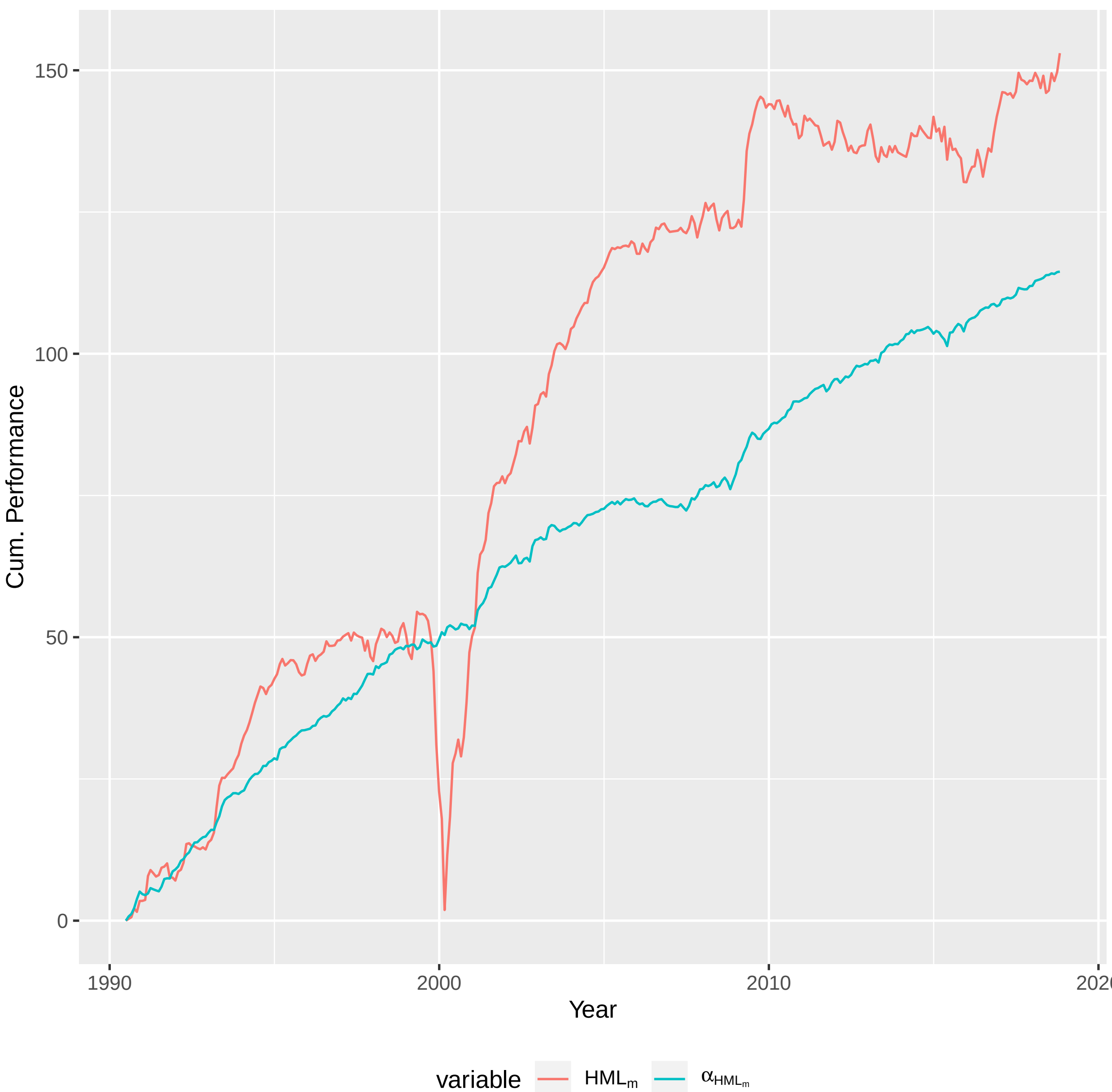

Table A.1: Constituent lists: Developed markets

The table contains the Research lists, Worldscope lists and Dead lists of developed markets countries in my sample.

Country	Lists	Country	Lists
Australia	DEADAU	**Italy**	DEADIT
	FAUS		FITA
	WSCOPEAU		WSCOPEIT
Austria	DEADOE	**Japan**	DEADJP
	FOST		FFUKUOKA
	WSCOPEOE		FJASDAQ
Belgium	DEADBG		FOSAKA
	FBEL		FTOKYO
	FBELAM		JAPOTC
	FBELCM		WSCOPEJP
	WSCOPEBG	**Netherlands**	DEADNL
Canada	DEADCN1		FHOL
	DEADCN2		WSCOPENL
	DEADCN3	**New Zealand**	DEADNZ
	DEADCN4		FNWZ
	DEADCN5		WSCOPENZ
	DEADCN6	**Norway**	DEADNW
	FTORO		FNOR
	FVANC		WSCOPENW
	LTTOCOMP	**Portugal**	DEADPT
	WSCOPECN		FPOR
Denmark	DEADDK		WSCOPEPT
	FDEN	**Singapore**	DEADSG
	WSCOPEDK		FSIN
Finland	DEADFN		FSINQ
	FFIN		WSCOPESG
	WSCOPEFN	**Spain**	DEADES
France	DEADFR		FSPN
	FFRA		WSCOPEES
	WSCOPEFR	**Sweden**	DEADSD
Germany	DEADBD1		FAKTSWD
	DEADBD2		FSWD
	DEADBD3		WSCOPESD
	DEADBD4	**Switzerland**	DEADSW
	DEADBD5		FSWA
	DEADBD6		FSWS
	FGER1		FSWUP
	FGER2		WSCOPESW
	FGERIBIS	**United Kingdom**	DEADUK
	FGKURS		FBRIT
	WSCOPEBD		LSETSCOS
Hong Kong	DEADHK		LSETSMM
	FHKQ		LUKPLUSM
	WSCOPEHK		WSCOPEJE
Ireland	DEADIR		WSCOPEUK
	FIRL		
	WSCOPEIR		
Israel	DEADIS		
	FISRAEL		
	WSCOPEIS		

Table A.2: Constituent lists: Emerging markets

The table contains the Research lists, Worldscope lists and Dead lists of emerging markets countries in my sample.

Argentina	DEADAR	**Pakistan**	DEADPA
	FPARGA		FPAK
	WSCOPEAR		FPAKUP
Brazil	DEADBRA		WSCOPEPK
	FBRA	**Peru**	DEADPE
	WSCOPEBR		FPERU
Chile	DEADCHI		WSCOPEPE
	FCHILE	**Philippines**	DEADPH
	FCHILE10		FPHI
	WSCOPECL		FPHILA
China	DEADCH		FPHIMN
	FCHINA		FPHIQ
	WSCOPECH		WSCOPEPH
Colombia	DEADCO	**Poland**	DEADPO
	FCOL		FPOL
	WSCOPECB		WSCOPEPO
Czech Republic	DEADCZ	**Qatar**	DEADQT
	FCZECH		FQATAR
	FCZECHUP		WSCOPEQA
	WSCOPECZ	**Russia**	DEADRU
Egypt	DEADEGY		FRTSCL
	EGYPTALL		FRUS
	FEGYPT		FRUSUP
	WSCOPEEY		WSCOPERS
Greece	DEADGR	**Slovakia**	ALLSLOV
	FGREE		DEADSLO
	FGRMM		FSLOVAK
	FGRPM		WSCOPESX
	FNEXA	**South Africa**	DEADSAF
	WSCOPEGR		FSAF
Hungary	DEADHU		WSCOPESA
	FHUN		DEADKO
	WSCOPEHN		FKONEX
India	DEADIND		FKOR
	FBSE		WSCOPEKO
	FINDIA	**Sri Lanka**	DEADSL
	FINDNW		FSRILA
	FINDUP		FSRIUP
	FNSE		WSCOPECY
	WSCOPEIN	**Taiwan**	DEADTW
Indonesia	DEADIDN		FTAIQ
	FINO		WSCOPETA
	WSCOPEID	**Thailand**	DEADTH
Jordan	DEADJO		FTHAQ
	FJORD		WSCOPETH
	WSCOPEJO	**Turkey**	DEADTK
Malaysia	DEADMY		FTURK
	FMAL		FTURKUP
	FMALQ		WSCOPETK
	WSCOPEMY	**United Arab Emirates**	DEADAB
Mexico	DEADME		DEADDB
	FMEX		FABUD
	MEX101		FDUBAI
	WSCOPEMX		WSCOPEAE
Morocco	DEADMOR	**Venezuela**	DEADVE
	FMOR		FVENZ
	WSCOPEMC		WSCOPEVE

Table A.3: Static Screens
The table displays the static screens applied in my study, mainly following Ince and Porter (2006), Schmidt et al. (2017) and Griffin et al. (2010). Column 3 lists the Datastream items involved (on the left of the equality sign) and the values which I set them to in the filter process (on the right of the equality sign). Column 4 indicates the source of the screens.

Nr.	Description	Datastream item(s) involved	Source
(1)	For firms with more than one security, only the one with the biggest market capitalization and liquidity is used.	MAJOR = Y	Schmidt et al. (2017)
(2)	The type of security must be equity.	TYPE = EQ	Ince and Porter (2006)
(3)	Only the primary quotations of a security are analyzed.	ISINID = P	Fong, Holden, and Trzcinka (2017)
(4)	Firms are located in the respective domestic country.	GEOGN = country shortcut	Ince and Porter (2006)
(5)	Securities are listed in the respective domestic country.	GEOLN = country shortcut	Griffin et al. (2010)
(6)	Securities with quoted currency different from the one of the associated country are disregarded.[a]	PCUR = currency shortcut of the country	Griffin et al. (2010)
(7)	Securities with ISIN country code different from the one of the associated country are disregarded.[b]	GGISN = country shortcut	Annaert, Ceuster, and Verstegen (2013)
(8)	Securities whose name fields indicate non-common stock affiliation are disregarded.	NAME, ENAME, ECNAME	Ince and Porter (2006), Campbell et al. (2010), Griffin et al. (2010) and Karolyi et al. (2012)

[a] In this filter rule also the respective pre-euro currencies are accepted for countries within the euro zone. Moreover, in Russia "USD" is also accepted as currency, besides "RUB".
[b] In Hong Kong, ISIN country codes equal to "BM" or "KY" and in the Czech Republic ISIN country codes equal to "CS" are also accepted.

Table A.4: Generic Keyword Deletions

The table reports the generic keywords, which are searched for in the names of all stocks of all countries. If a harmful keyword is detected as part of the name of a stock, the respective stock is removed from the sample.

Non-common equity	Keywords
Duplicates	1000DUPL, DULP, DUP, DUPE, DUPL, DUPLI, DUPLICATE, XSQ, XETa
Depository Receipts	ADR, GDR
Preferred Stock	PF, 'PF', PFD, PREF, PREFERRED, PRF
Warrants	WARR, WARRANT, WARRANTS, WARRT, WTS, WTS2
Debt	%, DB, DCB, DEB, DEBENTURE, DEBENTURES, DEBT
Unit Trusts	.IT, .ITb, TST, INVESTMENT TRUST, RLST IT, TRUST, TRUST UNIT, TRUST UNITS, TST, TST UNIT, TST UNITS, UNIT, UNIT TRUST, UNITS, UNT, UNT TST, UT
ETFs	AMUNDI, ETF, INAV, ISHARES, JUNGE, LYXOR, X-TR
Expired securities	EXPD, EXPIRED, EXPIRY, EXPY
Miscellaneous (mainly taken from Ince and Porter (2006))	ADS, BOND, CAP.SHS, CONV, DEFER, DEP, DEPY, ELKS, FD, FUND, GW.FD, HI.YIELD, HIGH INCOME, IDX, INC.&GROWTH, INC.&GW, INDEX, LP, MIPS, MITS, MITT, MPS, NIKKEI, NOTE, OPCVM, ORTF, PARTNER, PERQS, PFC, PFCL, PINES, PRTF, PTNS, PTSHP, QUIBS, QUIDS, RATE, RCPTS, REAL EST, RECEIPTS, REIT, RESPT, RETUR, RIGHTS, RST, RTN.INC, RTS, SBVTG, SCORE, SPDR, STRYPES, TOPRS, UTS, VCT, VTG.SAS, XXXXX, YIELD, YLD

Table A.5: Country-Specific Keyword Deletions

The table reports the country-specific keywords, which are searched for in the names of all stocks of the respective countries. If a harmful keyword is detected as part of the name of a stock, the respective stock is removed from the sample.

Country	Keywords
Australia	PART PAID, RTS DEF, DEF SETT, CDI
Austria	PC, PARTICIPATION CERTIFICATE, GENUSSSCHEINE, GENUSSCHEINE
Belgium	VVPR, CONVERSION, STRIP
Brazil	PN, PNA, PNB, PNC, PND, PNE, PNF, PNG, RCSA, RCTB
Canada	EXCHANGEABLE, SPLIT, SPLITSHARE, VTG\\., SBVTG\\., VOTING, SUB VTG, SERIES
Denmark	\\)CSE\\)
Finland	USE
France	ADP, CI, SICAV, \\)SICAV\\), SICAV-
Germany	GENUSSCHEINE
Greece	PR
India	FB DEAD, FOREIGN BOARD
Israel	P1, 1, 5
Italy	RNC, RP, PRIVILEGIES
Korea	1P
Mexico	'L', 'C'
Malaysia	'A'
Netherlands	CERTIFICATE, CERTIFICATES, CERTIFICATES\\), CERT, CERTS, STK\\.
New Zealand	RTS, RIGHTS
Peru	INVERSION, INVN, INV
Philippines	PDR
South Africa	N', OPTS\\., CPF\\., CUMULATIVE PREFERENCE
Sweden	CONVERTED INTO, USE, CONVERTED-, CONVERTED - SEE
Switzerland	CONVERTED INTO, CONVERSION, CONVERSION SEE
United Kingdom	PAID, CONVERSION TO, NON VOTING, CONVERSION 'A'

Table A.6: Dynamic Screens

The table displays the dynamic screens applied to the data in my study, following Ince and Porter (2006), Griffin et al. (2010), Griffin et al. (2011), Jacobs (2016) and Schmidt et al. (2017). Column 3 lists the respective Datastream items. Column 4 refers to the source of the screens.

Nr.	Description	Datastream item(s) involved	Source
(1)	I delete the zero returns at the end of the return time-series, which exist, because in case of a delisting Datastream displays stale prices from the date of delisting until the end of the respective time-series. I also delete the associated market capitalizations.	TRI, MV	Ince and Porter (2006)
(2)	I delete the associated returns and market capitalizations in case of abnormal prices (unadjusted prices > 1000000).	TRI, MV, UP	The screen originally stems from Schmidt et al. (2017), whereby I employ it on the unadjusted price.
(3)	I delete monthly returns and the associated market capitalizations in case of return spikes (returns > 990%).	TRI, MV	Schmidt et al. (2017)
(4)	I delete monthly returns and the associated market capitalizations in case of strong return reversals, defined as follows: R_{t-1} or $R_t >= 3.0$ and $(1 + R_{t-1})(1 + R_t) - 1 < 0.5$.	TRI, MV	Ince and Porter (2006)
(5)	I delete observations of stocks that show non-zero price changes in less than 50% of the traded months in previous 12 months.	TRI, MV	Griffin et al. (2011)

References

Ammann, M., Odoni, S., Oesch, D., 2012. An alternative three-factor model for international markets: Evidence from the European Monetary Union. Journal of Banking & Finance 36, 1857–1864.

Annaert, J., Ceuster, M. D., Verstegen, K., 2013. Are extreme returns priced in the stock market? European evidence. Journal of Banking & Finance 37, 3401–3411.

Asness, C., Frazzini, A., Israel, R., Moskowitz, T., 2015. Fact, fiction, and value investing. Journal of Portfolio Management 42, 34–52.

Asness, C. S., Frazzini, A., 2013. The devil in HML's details. Journal of Portfolio Management 39, 49–68.

Ball, R., Gerakos, J., Linnainmaa, J. T., Nikolaev, V. V., 2015. Deflating profitability. Journal of Financial Economics 117, 225–248.

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

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

Barillas, F., Kan, R., Robotti, C., Shanken, J., 2019. Model comparison with Sharpe ratios. Journal of Financial and Quantitative Analysis , forthcoming, 1–35.

Barillas, F., Shanken, J., 2017. Which alpha? Review of Financial Studies 30, 1316–1338.

Barillas, F., Shanken, J., 2018. Comparing asset pricing models. Journal of Finance 73, 715–754.

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.

Cakici, N., Fabozzi, F. J., Tan, S., 2013. Size, value, and momentum in emerging market stock returns. Emerging Markets Review 16, 46–65.

Campbell, C. J., Cowan, A. R., Salotti, V., 2010. Multi-country event-study methods. Journal of Banking & Finance 34, 3078–3090.

Carhart, M. M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57–82.

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

Daniel, K., Hirshleifer, D., Sun, L., 2019. Short- and long-horizon behavioral factors. Review of Financial Studies , forthcoming.

Daniel, K., Titman, S., 2006. Market reactions to tangible and intangible information. Journal of Finance 61, 1605–1643.

de Groot, W., Pang, J., Swinkels, L., 2012. The cross-section of stock returns in frontier emerging markets. Journal of Empirical Finance 19, 796–818.

Fama, E. F., French, K. R., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.

Fama, E. F., French, K. R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Fama, E. F., French, K. R., 1998. Value versus growth: The international evidence. Journal of Finance 53, 1975–1999.

Fama, E. F., French, K. R., 2012. Size, value, and momentum in international stock returns. Journal of Financial Economics 105, 457–472.

Fama, E. F., French, K. R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1 – 22.

Fama, E. F., French, K. R., 2016. Dissecting anomalies with a five-factor model. Review of Financial Studies 29, 69–103.

Fama, E. F., French, K. R., 2017. International tests of a five-factor asset pricing model. Journal of Financial Economics 123, 441–463.

Fama, E. F., French, K. R., 2018. Choosing factors. Journal of Financial Economics 128, 234–252.

Fletcher, J., 2018. Bayesian tests of global factor models. Journal of Empirical Finance 48, 279–289.

Fletcher, J., 2019. Model comparison tests of linear factor models in U.K. stock returns. Finance Research Letters 28, 281–291.

Fong, K. Y. L., Holden, C. W., Trzcinka, C. A., 2017. What are the best liquidity proxies for global research? Review of Finance 21, 1355–1401.

Gibbons, M. R., Ross, S. A., Shanken, J., 1989. A test of the efficiency of a given portfolio. Econometrica 57, 1121–1152.

Goyal, A., Wahal, S., 2015. Is momentum an echo? Journal of Financial and Quantitative Analysis 50, 1237–1267.

Green, J., Hand, J. R. M., Zhang, X. F., 2017. The characteristics that provide independent information about average U.S. monthly stock returns. Review of Financial Studies 30, 4389–4436.

Griffin, J. M., 2002. Are the Fama and French factors global or country specific? Review of Financial Studies 15, 783–803.

Griffin, J. M., Hirschey, N. H., Kelly, P. J., 2011. How important is the financial media in global markets? Review of Financial Studies 24, 3941–3992.

Griffin, J. M., Kelly, P. J., Nardari, F., 2010. Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets. Review of Financial Studies 23, 3225–3277.

Hanauer, M. X., Huber, D., 2019. Constructing a powerful profitability factor: International evidence. SSRN Working Paper no. 3234436 .

Hanauer, M. X., Lauterbach, J. G., 2019. The cross-section of emerging market stock returns. Emerging Markets Review 38, 265–286.

Harvey, C. R., 2017. Presidential address: The scientific outlook in financial economics. Journal of Finance 72, 1399–1440.

Harvey, C. R., Liu, Y., Zhu, H., 2016. … and the cross-section of expected returns. Review of Financial Studies 29, 5–68.

Heston, S. L., Sadka, R., 2010. Seasonality in the cross section of stock returns: The international evidence. Journal of Financial and Quantitative Analysis 45, 1133–1160.

Hirshleifer, D., Hou, K., Teoh, S. H., Zhang, Y., 2004. Do investors overvalue firms with bloated balance sheets? Journal of Accounting and Economics 38, 297–331.

Hirshleifer, D., Hsu, P.-H., Li, D., 2013. Innovative efficiency and stock returns. Journal of Financial Economics 107, 632–654.

Hou, K., Karolyi, G. A., Kho, B.-C., 2011. What factors drive global stock returns? Review of Financial Studies 24, 2527–2574.

Hou, K., Mo, H., Xue, C., Zhang, L., 2019. Which factors? Review of Finance 23, 1–35.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. Review of Financial Studies 28, 650–705.

Hou, K., Xue, C., Zhang, L., 2018. Replicating anomalies. Review of Financial Studies , forthcoming.

Ince, O. S., Porter, R. B., 2006. Individual equity return data from Thomson Datastream: Handle with care! Journal of Financial Research 29, 463–479.

Jacobs, H., 2016. Market maturity and mispricing. Journal of Financial Economics 122, 270–287.

Jacobs, H., Müller, S., 2020. Anomalies across the globe: Once public, no longer existent? Journal of Financial Economics 135, 213–230.

Jacobs, H., Müller, S., 2018. ... and nothing else matters? On the dimensionality and predictability of international stock returns. SSRN Working Paper no. 2845306 .

Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. Journal of Finance 45, 881–898.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. Journal of Finance 48, 65–91.

Kan, R., Wang, X., Zheng, X., 2019. In-sample and out-of-sample sharpe ratios of multi-factor asset pricing models. SSRN Working Paper no. 3454628 .

Karolyi, A. G., 2016. Home bias, an academic puzzle. Review of Finance 20, 2049–2078.

Karolyi, G. A., Lee, K.-H., van Dijk, M. A., 2012. Understanding commonality in liquidity around the world. Journal of Financial Economics 105, 82–112.

Keloharju, M., Linnainmaa, J. T., Nyberg, P. M., 2020. Are return seasonalities due to risk or mispricing? Evidence from seasonal reversals. Journal of Financial Economics , forthcoming.

Lev, B., Srivastava, A., 2019. Explaining the recent failure of value investing. SSRN Working Paper no. 3442539 .

Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Review of Economics and Statistics 47, 13–37.

Lu, X., Stambaugh, R. F., Yuan, Y., 2017. Anomalies abroad: Beyond data mining. NBER Working Paper No. 23809 .

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

McLean, R. D., Pontiff, J., Watanabe, A., 2009. Share issuance and cross-sectional returns: International evidence. Journal of Financial Economics 94, 1–17.

Novy-Marx, R., 2013. The other side of value: The gross profitability premium. Journal of Financial Economics 108, 1–28.

Novy-Marx, R., 2015. How can a q-theoretic model price momentum? NBER Working Paper No. 20985 .

Park, H., 2019. An intangible-adjusted book-to-market ratio still predicts stock returns. Critical Finance Review , forthcoming.

Pontiff, J., Woodgate, A., 2008. Share issuance and cross-sectional returns. Journal of Finance 63, 921–945.

Rosenberg, B., Reid, K., Lanstein, R., 1985. Persuasive evidence of market inefficiency. Journal of Portfolio Management 11, 9–16.

Rouwenhorst, K. G., 1998. International momentum strategies. Journal of Finance 53, 267–284.

Schmidt, P. S., Von Arx, U., Schrimpf, A., Wagner, A. F., Ziegler, A., 2017. On the construction of common size, value and momentum factors in international stock markets: A guide with applications. Swiss Finance Institute Research Paper 10.

Sharpe, W. F., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. Journal of Finance 19, 425–442.

Sloan, R. G., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? The Accounting Review 71, 289–315.

Stambaugh, R. F., Yuan, Y., 2017. Mispricing factors. Review of Financial Studies 30, 1270–1315.

Thomson Reuters, 2013. Worldscope database datatype definitions guide. Issue 14.2 ed.

Walkshäusl, C., Lobe, S., 2014. The alternative three-factor model: An alternative beyond US markets? European Financial Management 20, 33–70.

Walkshäusl, C., Lobe, S., 2015. The enterprise multiple investment strategy: International evidence. Journal of Financial and Quantitative Analysis 50, 781–800.

Watanabe, A., Xu, Y., Yao, T., Yu, T., 2013. The asset growth effect: Insights from international equity markets. Journal of Financial Economics 108, 529–563.

Zaremba, A., Czapkiewicz, A., 2017. Digesting anomalies in emerging European markets: A comparison of factor pricing models. Emerging Markets Review 31, 1–15.

Zaremba, A., Maydybura, A., 2019. The cross-section of returns in frontier equity markets: Integrated or segmented pricing? Emerging Markets Review 38, 219–238.