Factor Zoo (.zip)*

Alexander Swade[1,2], Matthias X. Hanauer[2,3], Harald Lohre[1,2] and David Blitz[2]

[1]Lancaster University Management School †
[2]Robeco ‡
[3]Technical University of Munich §

This version: November 15, 2023

Abstract

The number of factors allegedly driving the cross-section of stock returns has grown steadily over time. We explore how much this 'factor zoo' can be compressed, focusing on explaining the available alpha rather than the covariance matrix of factor returns. Our findings indicate that about 15 factors are enough to span the entire factor zoo. This evidence suggests that many factors are redundant but also that merely using a handful of factors, as in common asset pricing models, is insufficient. While the selected factor styles remain persistent, the specific style representatives vary over time, underscoring the importance of continuous factor innovation.

Keywords: Factor zoo, factor model, factor investing, alpha, GRS test

JEL Classification: G12, G14, G15

*Note that this paper expresses the authors' views which do not necessarily coincide with those of Robeco. We thank Amit Goyal, Clint Howard, Stefan Mittnik, Michael Rockinger, Pim van Vliet, and participants at the 2023 CEQURA Conference on Advances in Financial and Insurance Risk Management in Munich and the Robeco Research Seminar in Rotterdam for helpful comments and suggestions. This work has been supported by an ESRC NWSSDTP CASE Grant. Corresponding author: Alexander Swade, a.swade@lancaster.ac.uk

†Lancaster University Management School, Bailrigg, Lancaster LA1 4YX, United Kingdom.

‡Robeco Institutional Asset Management B.V., Weena 850, 3014 DA Rotterdam, the Netherlands.

§Technical University of Munich, Arcisstraße 21, 80333 Munich, Germany.

Factor Zoo (.zip)

This version: November 15, 2023

Abstract

The number of factors allegedly driving the cross-section of stock returns has grown steadily over time. We explore how much this 'factor zoo' can be compressed, focusing on explaining the available alpha rather than the covariance matrix of factor returns. Our findings indicate that about 15 factors are enough to span the entire factor zoo. This evidence suggests that many factors are redundant but also that merely using a handful of factors, as in common asset pricing models, is insufficient. While the selected factor styles remain persistent, the specific style representatives vary over time, underscoring the importance of continuous factor innovation.

Keywords: Factor zoo, factor model, factor investing, alpha, GRS test

JEL Classification: G12, G14, G15

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

However, over the last 25 years, hundreds of factors have emerged in the literature, all of which allegedly offer a unique new source of return. Cochrane (2011) aptly characterizes the state of play as a 'zoo of factors' that needs to be tamed and structured. As the existing factor models of Fama and French (1993) and Carhart (1997) cannot explain many of the new factors, Fama and French (2015, 2018) extend these models to five- and six-factor models by adding investment (Cooper, Gulen, and Schill, 2008) and profitability (Novy-Marx, 2013) factors. These factor models compete with alternative four-factor models such as the Hou, Xue, and Zhang (2015) q-factor model and the Stambaugh and Yuan (2017) mispricing model or the revised six-factor model of Barillas et al. (2020). Although these models use different factors, there seems to be a consensus among leading academics that most of the factor zoo can be explained by parsimonious models consisting of just four to six factors, see Bartram et al. (2021).

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

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

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

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

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

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

The remainder of this paper is structured as follows. Section 2 outlines our method for identifying the most important factors in the factor zoo. Section 3 presents our empirical results for the U.S. factor zoo, taking into account different weighting schemes and dynamic time periods. Next, we analyze the global factor zoo in Section 4 and test the sensitivity of our factor selection method to different regions. Section 5 concludes.

2 Methodology

2.1 Identifying factors that compress the factor zoo

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

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

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

(1998), and the approach is applied for model comparisons in Barillas et al. (2020) and Hanauer (2020).

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

Factor selection steps

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

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

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

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

$$f_i = \alpha_i + \beta_m r_m + \sum_{k=1}^{l} \beta_k f_k + \beta^{test} f^{test} + \varepsilon_i \qquad i = 1, \ldots, N - l \tag{2}$$

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

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

$$n(\alpha)_{t>x} = |\{a_i \,|\, t(a_i) > x\}| \qquad i = 1, \ldots, N - l \tag{3}$$

where x is the selected significance threshold.

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

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

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

as well as a more conservative one where $t > 3.00$.

2.2 Evaluating factor models

When looking to span the whole factor zoo one is dealing with nested models. A common metric in this field is the GRS statistic of Gibbons, Ross, and Shanken (1989), which produces a test of whether candidate factors help to improve a given model's explanation of expected returns. Specifically, the GRS test investigates whether the alphas of the test assets are jointly different from zero. The GRS test is widely used in empirical finance and has become a standard tool for evaluating the performance of asset pricing models as, e.g., in Fama and French (1996, 2015) and Stambaugh and Yuan (2017). Formally, the empirical GRS statistic is given as follows: Consider an asset pricing model (2) with K factors, N test assets, and τ return observations for each time-series. We follow Fama and French (2018) and define the maximum squared Sharpe ratio for the intercepts as

$$Sh^2(\alpha) = \alpha^\intercal \Sigma^{-1} \alpha \tag{4}$$

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

$$Sh^2(f) = \overline{f}^\intercal \Omega^{-1} \overline{f} \tag{5}$$

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

$$F_{GRS} = \frac{\tau(\tau - N - K)}{N(\tau - K - 1)} \frac{Sh^2(\alpha)}{(1 + Sh^2(f))} \tag{6}$$

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

Note that the GRS statistic is crucially determined by the ratio of $Sh^2(\alpha)$ and $Sh^2(f)$. When evaluating factor models, the goal is to identify a model that observes the smallest maximum squared Sharpe ratios for the alphas, and thus captures most of the return variation through its systematic components. Given the relation in equation (6), Barillas and Shanken (2017) propose to use the factors' maximum squared Sharpe ratio to evaluate the power of a set of candidate models. Fama and French (2018) further analyze the resulting implications and conclude that the model reducing $Sh^2(\alpha)$ the most is also the model with the highest $Sh^2(f)$, consistent with Barillas and Shanken (2017). Thus, our empirical analysis will not only report GRS statistics and their associated p-values but also $Sh^2(f)$s and average absolute alphas, labeled $Avg|\alpha|$.

3 Compressing the factor zoo

3.1 Data

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

To enable covering all 153 factors in our study, we start our investigation of U.S. capped

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

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

value-weighted factors in November 1971, and our sample period ends in December 2021. In capped value-weighted factors, stocks are sorted into characteristic terciles each month and the capped value-weighted tercile returns are calculated as the market equity-weighted portfolio returns capped at the NYSE 80th percentile. The factor return is then defined as the high- minus low-tercile return, cf. JKP. This factor construction is designed to create tradable yet balanced portfolios that are neither dominated by mega nor tiny caps. Note that we check for robustness of our main analysis with respect to alternative factor weighting schemes in Section 3.5.

Figure 1 provides an overview of all factors' annualized alphas based on monthly CAPM regressions. All factors are clustered into 13 categories as identified by JKP based on hierarchical agglomerative clustering (Murtagh and Legendre, 2014), and cluster names are driven by the most representative characteristics. We observe mostly positive annualized alphas for all clusters but the Low Leverage cluster. The average alpha across all factors is 3.51% p.a., and alphas are fairly evenly distributed across and within clusters.

[Figure 1 about here.]

It is not surprising that most of the factors exhibit a significant alpha premium, i.e., indicating incremental power beyond the market return, and thus are deemed to be a relevant factor to begin with. Blitz (2023) further analyzes this set of factors regarding their market risk, performance cyclicality, and inherent seasonal and momentum effects in the cross-section of factor returns. Yet, the question of the incremental alpha contribution of individual factors within the factor zoo is an open question.

3.2 Main results

Table 1 depicts our main results following the iterative factor selection process described in Section 2. We report the selected factors, their associated factor style cluster, the GRS statistics and corresponding p-values, $p(GRS)$, the average absolute intercept $Avg|a|$, the maximum squared Sharpe ratio for the model's factors $Sh^2(f)$, as well as their Sharpe ratios, SR. Columns 10–11 refer to the number of significant factor' alphas after controlling for the specified factor model and thus indicate the incremental explanatory power of the selected factor. The column labels $t > 2$ and $t > 3$ refer to our iterative factor selection with a significance alpha threshold of $t(\alpha) > 1.96$ or $t(\alpha) > 3.00$, respectively. Note that we also report the number of significant factors under common factor models in the subsequent table.

[Table 1 about here.]

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

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

Iterating further, the factor model increases by construction. Whilst Table 1 documents the impact of adding the thirty most relevant factors, it only takes half of that number to span the whole factor zoo. Adding the 15th factor (*highest five days of return scaled by volatility*, $rmax5\_rvol\_21d$) the number of remaining significant alphas drops to zero ($t > 3$). Even with the less strict threshold of $t > 2$, it only takes a total of 18 iterations to render the remaining alphas insignificant. These cut-off numbers are in line with the alternative stopping criterion based on the significance level of the GRS statistic: *highest five days of return scaled by volatility* ($rmax5\_rvol\_21d$) is also the first factor exceeding the 5% significance level with a p-value of 0.09. Overall, these results indicate that it only takes 15 to 18 additional factors to span the factor zoo from an alpha perspective, regardless of the stopping criterion.

Against this backdrop, we wonder how our iterative factor selection compares to classic academic factor models. Table 2 reports the number of significant factor alphas under well-known models (measured at a significance threshold of $t > 3$). Here, columns FF5 and FF6 refer to the Fama and French (2015) five-factor model, where FF6 augments the latter by a momentum factor. Furthermore, HXZ, BS, and SY refer to the Hou, Xue, and Zhang (2015) q-factor model, the Barillas et al. (2020) revised six-factor model, and the Stambaugh and Yuan (2017) mispricing model, respectively.[3]

[Table 2 about here.]

Relative to the CAPM (86 alphas with $t > 3.00$), the Fama and French five and six-factor models, the q-factor model, and the mispricing model still leave between 58 and 69 alphas significant. However, the revised six-factor model of Barillas et al. (2020) substantially reduces the number of significant alphas to 33. This superior explanatory power stems from one main difference with the other models, namely the inclusion of the *cash-based operating profits-to-book assets* (cop_at) factor, which also emerged as a key factor in our iterative factor selection approach. The power of the iterative factor model approach reveals when we compare the academic factor models to the iterative factor models that contain the same number of factors. While the iterative model with four factors merely leaves ten significant

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

factors, the four-factor models of Hou, Xue, and Zhang (2015) and Stambaugh and Yuan (2017) still leave 60 and 55 significant alphas, respectively. Similarly, 14 and 15 remain significant for the iterative model with five and six factors, while for the Fama and French five and six-factor models and the Barillas et al. (2020) revised six-factor model 65, 53, and 29 alphas remain significant, respectively. These results reinforce that the selected factors do carry information over and above classic academic factors. More specifically, *cash-based operating profits-to-book assets* (cop_at) is the only factor from the 15 selected factors that is also contained in one of the common academic factor models, namely the Barillas et al. (2020) revised six-factor model. All the other selected factors represent either alternative value, profitability, investment, or momentum definitions or stem from alternative factor style clusters such as seasonality or short-term reversal that offer alpha beyond common factor models (cf., Blitz et al., 2023).

Note that the 15 selected factors emerge from 8 out of the 13 defined factor style categories and no factor from the remaining five categories is considered, see the highlighted factor bars in Figure 2. Moreover, the selected factors are not necessarily those with the highest CAPM alpha in a given factor style cluster; in fact, this only applies to the value, quality, short-term reversal, and seasonality clusters. Notably, whilst five of the eight represented factor clusters merely feature a single factor, the value, low risk, and investment clusters are represented by 3 to 4 factors.

[Figure 2 about here.]

Against this backdrop, we wonder how important it is to go with these selected factors or whether it is sufficient to determine the strongest factor from each of the 13 categories. The last column in Table 2 reports results for a 13-factor model consisting of the strongest (largest absolute CAPM alpha) factor per cluster. This model virtually spans the whole zoo, leaving just four factors unexplained. Out of the first 30 factors, it only fails to explain away the alpha from *sales growth for one quarter* ($saleq\_gr1$) and *intrinsic value-to-market*

($ival\_me$), highlighting the power of a cluster spanning model. Nevertheless, the iterative model with 13 additional factors only leaves one alpha significant.

3.3 The relevance of factors through time

Having analyzed the full sample evidence in 3.2, we wonder about the persistence of the individual factors' relevance through time. One caveat when analyzing the relative strengths of individual factors in the factor zoo is that many factors have only been published along the way and show a weaker post-publication performance (McLean and Pontiff, 2016). To better understand how the set of selected factors evolves through time, we repeat the iterative factor selection but restrict ourselves to the available factors at any point in time. Indeed, many factors that we find to work very well over the whole sample period were not known for many years. For example, the *residual momentum factor* ($resff3\_12\_1$) of Blitz, Huij, and Martens (2011) was only published in 2011 and would, therefore, not have been viable in the first 25 years of the sample. Specifically, our analysis is based on an expanding window analysis using an initial window of 180 months such that we obtain the first out-of-sample observation in December 1986. Each year, we only consider the already published factors when annually running the iterative factor selection and collect the information as presented in Table 1. However, note that we stop the iteration at the first occurrence of $n(\alpha)_{t(\alpha)>3} = 0$, i.e., when we arrive at the first factor model that renders all remaining factor alphas insignificant at a threshold of $t > 3.00$.

Figure 3 highlights the relevant factors through time, colored by their corresponding factor style cluster. That is, whenever a factor is chosen in the corresponding year's factor model, it is highlighted on the timeline. While the vast majority of factors are either never or rarely included, the top factors from the full sample evidence in Table 1 show up prominently, especially over the last 10–15 years.

[Figure 3 about here.]

We observe many factor style clusters to be included in the model for most of the time once a representative factor is published. For instance, the value cluster is present most of the time, but also the momentum cluster is constantly represented since its publication in Jegadeesh and Titman (1993). Other persistent factor categories are accruals, investment, seasonality, and short-term reversal.

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

3.4 Rolling window analysis

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

Figure 4 depicts the development of the iterative selection process, aggregated at the factor style cluster level. For each year in the rolling window analysis, we collect the selected factors by style factor cluster. The upper panel of Figure 4 plots the number of factors included in each year's model using a cut-off of $t > 2$, whilst the lower panel reports the results for a cut-off of $t > 3$. We report both thresholds to address data-mining concerns while simultaneously gauging the relevance of borderline factors. Also, using a window size of only 180 months naturally raises the bar for factors to exceed a given threshold, relative to the full sample evidence.

[Figure 4 about here.]

The upper panel of Figure 4 documents a decrease in the number of selected factors over time. Whilst it took some 15 factors to span the factor zoo in the early years of the sample period, this number decreased to about 8 factors in more recent years. Yet, there are specific factor styles that are persistently relevant through time, including low volatility, seasonality, investment, and quality. While momentum, short-term reversal, and value were included almost every year until the early 2010s, their relevance has though weakened towards the end of the sample.

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

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

3.5 Robustness regarding alternative weighting schemes

The weighting scheme used to construct factor portfolios can have a big impact, see Bessembinder, Burt, and Hrdlicka (2022) and Soebhag, Van Vliet, and Verwijmeren (2023) amongst others. We thus check for robustness of our results with respect to the three weighting schemes: capped value-weighting (CW), value-weighting (VW), and equal-weighting (EW). We repeat the analysis of Table 1 for the different weighting schemes and summarise the corresponding outcome in Figure 5.

[Figure 5 about here.]

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

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

In the last row of Figure 5, we track the number of remaining significant factors for a given factor model. The vertical dashed lines indicate the first occurrence of zero remaining significant factors for a given weighting scheme. Whilst all three weighting schemes reduce the number of remaining significant alphas with increasing size, the EW factor models are sometimes choppy in doing so. That is, increasing the number of factors does not necessarily lead to a decrease in the remaining significant factors in the zoo. Also, it takes 18 and 19 factors for CW and VW factor models to explain away all factor zoo alphas, while EW factor models would take more than 30 factors at a significance threshold of $t > 2$. Notably, this order is almost reversed at a threshold of $t > 3$. Then, VW factor models require 9 additional factors to eliminate significant alphas, followed by 11 and 15 factors for EW and CW, respectively. These results indicate that the selected CW factors are relatively stronger in explaining the CW factor zoo, whereas EW factor models are based on fewer strong factors, and the remainder are more subject to data mining concerns.

4 International evidence

4.1 Global factor selection

We next broaden our view and investigate whether the U.S. evidence from Table 1 carries over to global factors, using international data for 93 different countries. Given the limited availability of some stock-specific measures, we shorten our international sample and focus on the period from August 1993 to December 2021. The sample covers 136 common factors for the three regions World, U.S., and World ex U.S.[4]

Table 3 documents the iterative factor selection based on global factor data. Notwithstanding the use of global factors and a shorter sample period, we observe a good overlap in the selected factors compared to the U.S. results in Table 1. Out of the first ten selected factors, three are identical ($cop\_at, resff3\_12\_1, cowc\_gr1a$), and the two selected investment factors are close cousins of the U.S. ones. Also, the selected factor style cluster order is very similar (almost identical) for these top factors.

[Table 3 about here.]

The selection process reveals 11 global factors to span the global factor zoo when enforcing

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

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

However, the GRS statistics for the U.S. and World ex U.S. samples also decrease almost monotonically, indicating that the global factor models also work for these subsamples. Whilst the models in the U.S. sample have generally lower GRS statistics, resulting in a rejection of the null at about six factors, the models for the rest of the world experience higher GRS values. Note that the shorter sample period of this international analysis already induces a reduction in relevant factors by construction, cf. Figure 6 in the next subsection. Although the factor models derived from global data have explanatory power for the World ex U.S. factors, they do a better job on U.S. factors. This observation is likely explained by the much higher alphas that have to be explained as indicated by a GRS statistic of more than 7.03 for international factors compared to a GRS statistic of 2.13 for the U.S.

4.2 Regional comparisons

So far, global factors have proven relevant for explaining U.S. factor returns, but they lack explanatory power for the World ex US factors. Against the backdrop of Section 4.1, we wonder whether local factor models are indeed stronger than global ones. We thus determine the iterative factor model within each region separately and juxtapose the relevant statistics in Figure 6 by regions.

[Figure 6 about here.]

The GRS statistics are declining for all three regions by design. However, we clearly see that the World ex-U.S. factors' decline in GRS statistic occurs at a higher level than that of the U.S. and the global factors. As a result, it takes more than 30 factors to reject the null at a significance level of 5% for the World ex U.S. sample whilst U.S. and global models only take 15 and 22 factors, respectively. The differences are primarily driven by the slower convergence of the adjusted squared Sharpe ratios (*adj.* $Sh^2(f)$) in the different regions as shown in the second row of Figure 6. Whilst the U.S. factor models seem to converge to some limit in *adj.* $Sh^2(f)$, the other two models' respective lines still have a positive slope at 30 factors. Conversely, the average absolute alphas seem to converge to a statistically insignificant number for all three regions once 30 factors are considered.

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

5 Conclusion

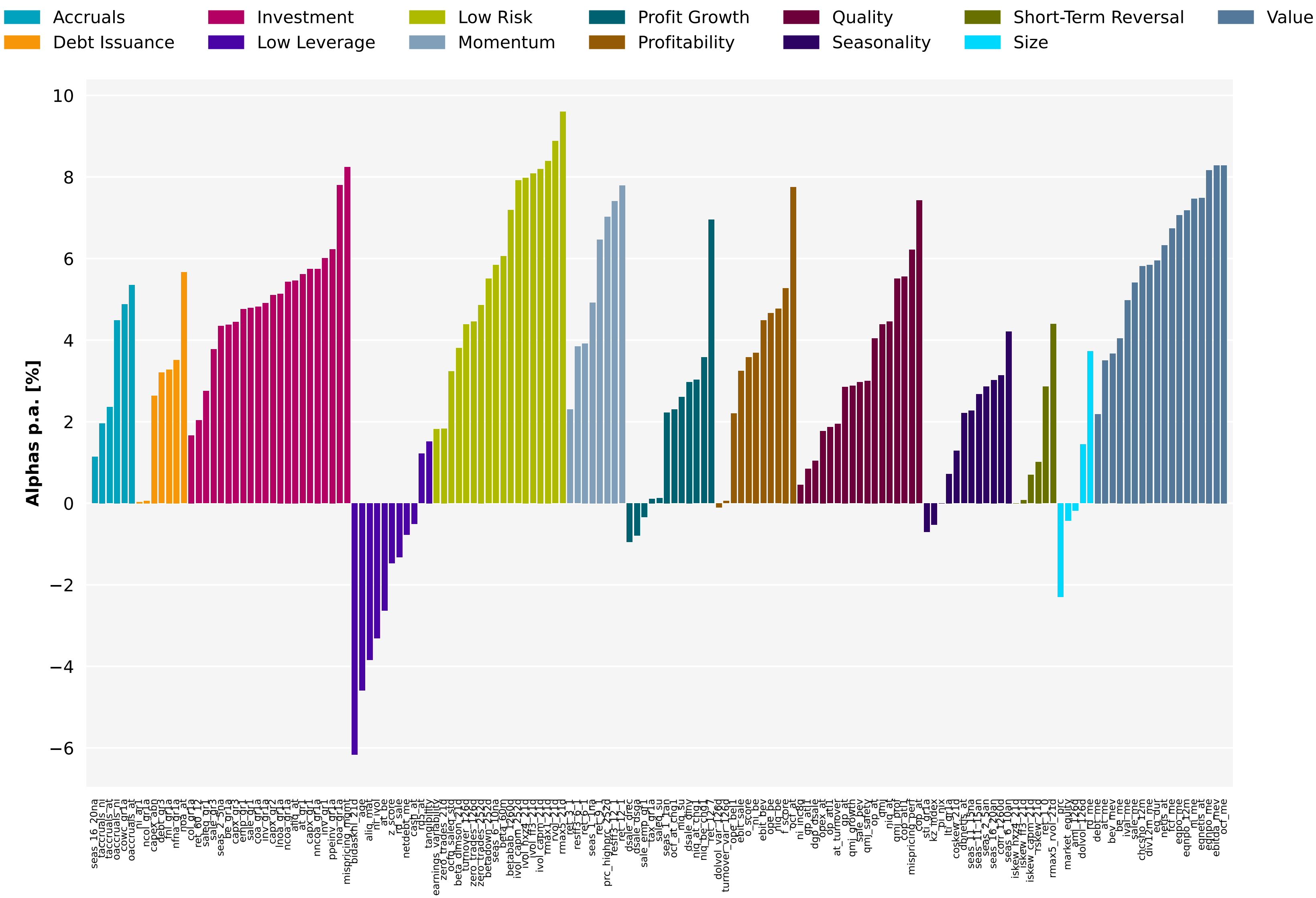
The factor zoo has grown significantly over the last decades as outlined in Cochrane (2011) and Harvey, Liu, and Zhu (2016), highlighting the need to separate sheep from goat factors. To this end, our study investigates the alpha contribution of individual factors in the factor zoo. Specifically, we propose an iterative factor selection strategy to compress the factor zoo, thus substantially reducing the number of factors without losing vital information for the tangency portfolio. The resulting factor sets capture the available alpha in the factor zoo with the minimum number of factors, which is relevant for both practitioners and academic purposes.

Using a comprehensive set of 153 U.S. equity factors, we find that a set of 10 to 20 factors spans the entire factor zoo, depending on the selected statistical significance level. This implies that most candidate factors are redundant but also that academic factor models, which typically contain just three to six factors, are too narrowly defined. When repeating the factor selection to factors as they become available over an expanding window, we find that newly published factors sometimes supersede older factor definitions, emphasizing the relevance of continuous factor innovation based on new insights or newly available data. However, the identified factor style clusters are quite persistent, emphasizing the relevance of diversification across factor styles.

Furthermore, we document that using equal-weighted factors as opposed to capped value-weighted factors requires more than 30 factors to span the factor zoo, indicating that equal-weighted factors exhibit stronger and more diverse alphas. Finally, applying our factor selection strategy to a set of global factors results in a similar set of selected factors. Although the factor models selected based on global data shrink the alpha for U.S. and World ex U.S. sub-universes, they perform better for the U.S., implying that international factors exhibit larger and more diverse alpha.

Overall, the proposed method effectively captures the available alpha in the factor zoo across different regions and subperiods, helping investors focus on the most relevant factors and providing academics with inspiration for enhancing asset pricing models.

Figure 1: Factor alphas



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

Table 1: Iterative factor selection

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

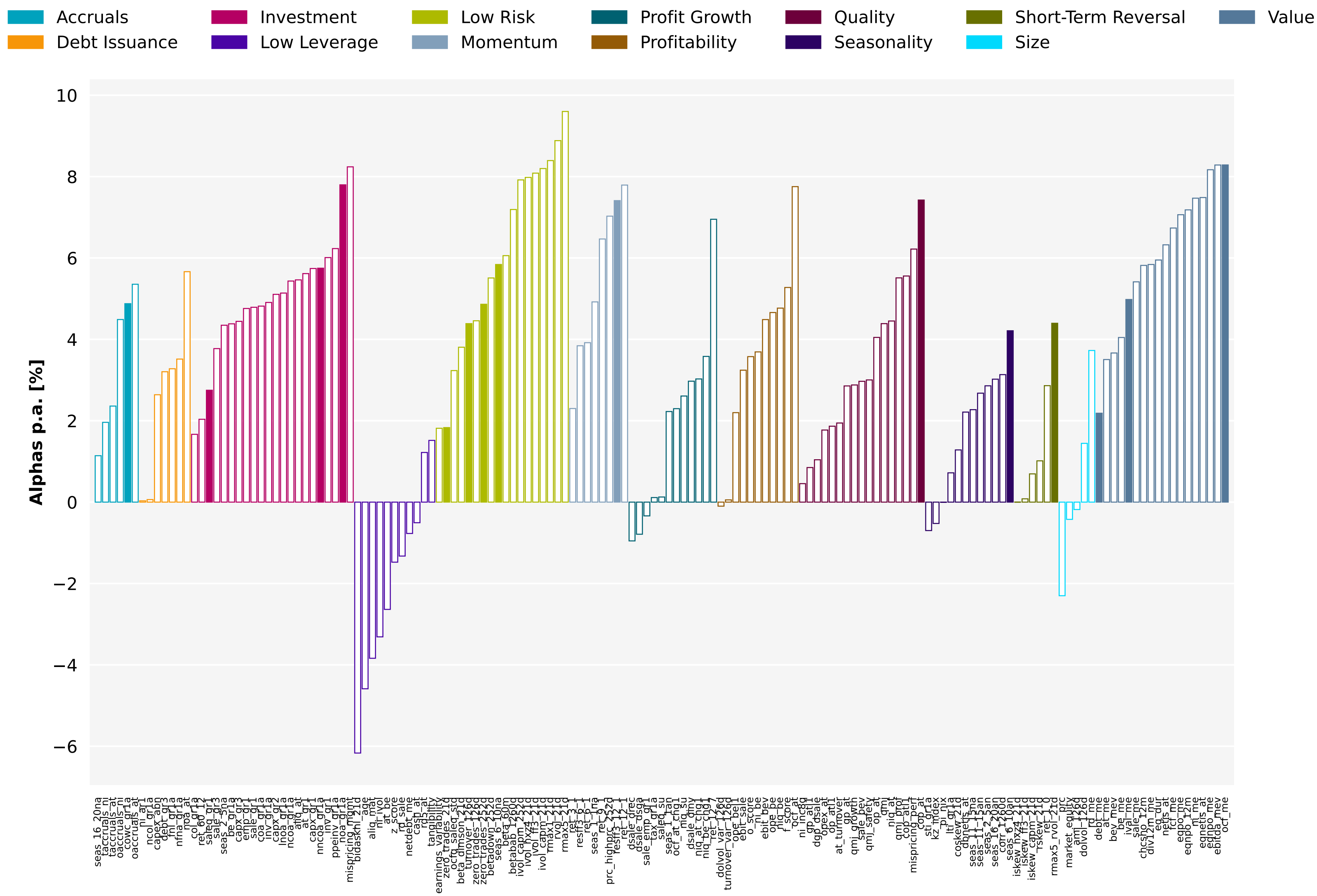
This table reports the results for an iterative factor model construction where the k-th iteration augments the model by the factor in row k. It shows the GRS statistic of Gibbons, Ross, and Shanken (1989) and its p-value, p(GRS); the annualised average absolute intercept $Avg|\alpha|$ in percentage,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

Table 2: Factor relevance in alternative models

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

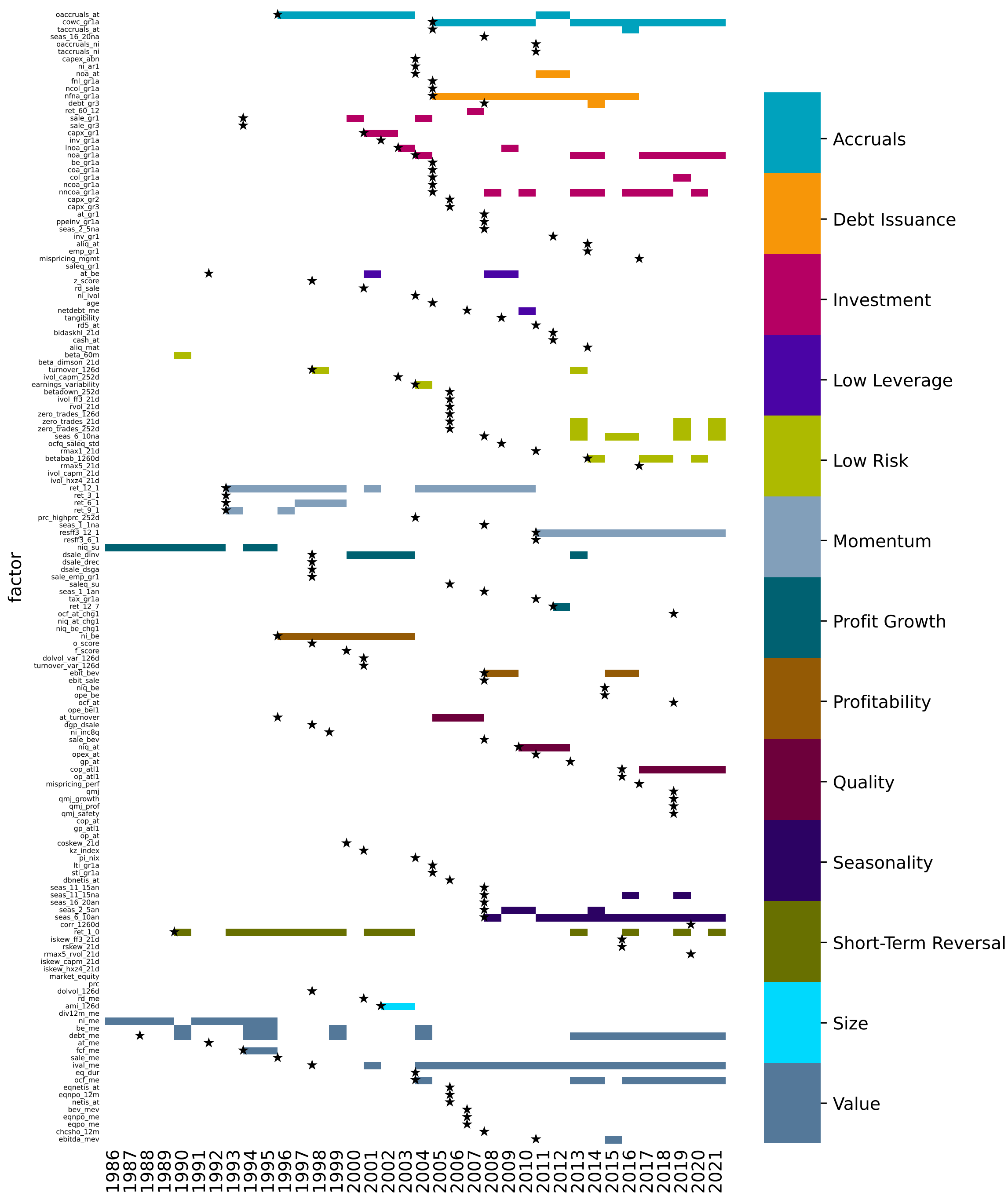
This table reports the results for an iterative factor model construction where the k-th iteration augments the model by the factor in row k. Columns 4-10 refer to the number of remaining significant factor alphas after controlling for the specified factor model. $t > 2$ and $t > 3$ control the factor zoo based on the iterative model using a significance alpha threshold of $t(\alpha) > 1.96$ and $t(\alpha) > 3.00$, respectively. FF5 refers to the Fama and French (2015) five-factor model whilst the augmented version include an additional momentum factor (FF6). HXZ refers to the model of Hou, Xue, and Zhang (2015), whilst BS and SY refer to the models of Barillas et al. (2020) and Stambaugh and Yuan (2017), respectively. C13 is based on the 13 strongest factors in the considered factor style clusters. The sample period is November 1971 to December 2021.

Figure 2: Selected alpha factors



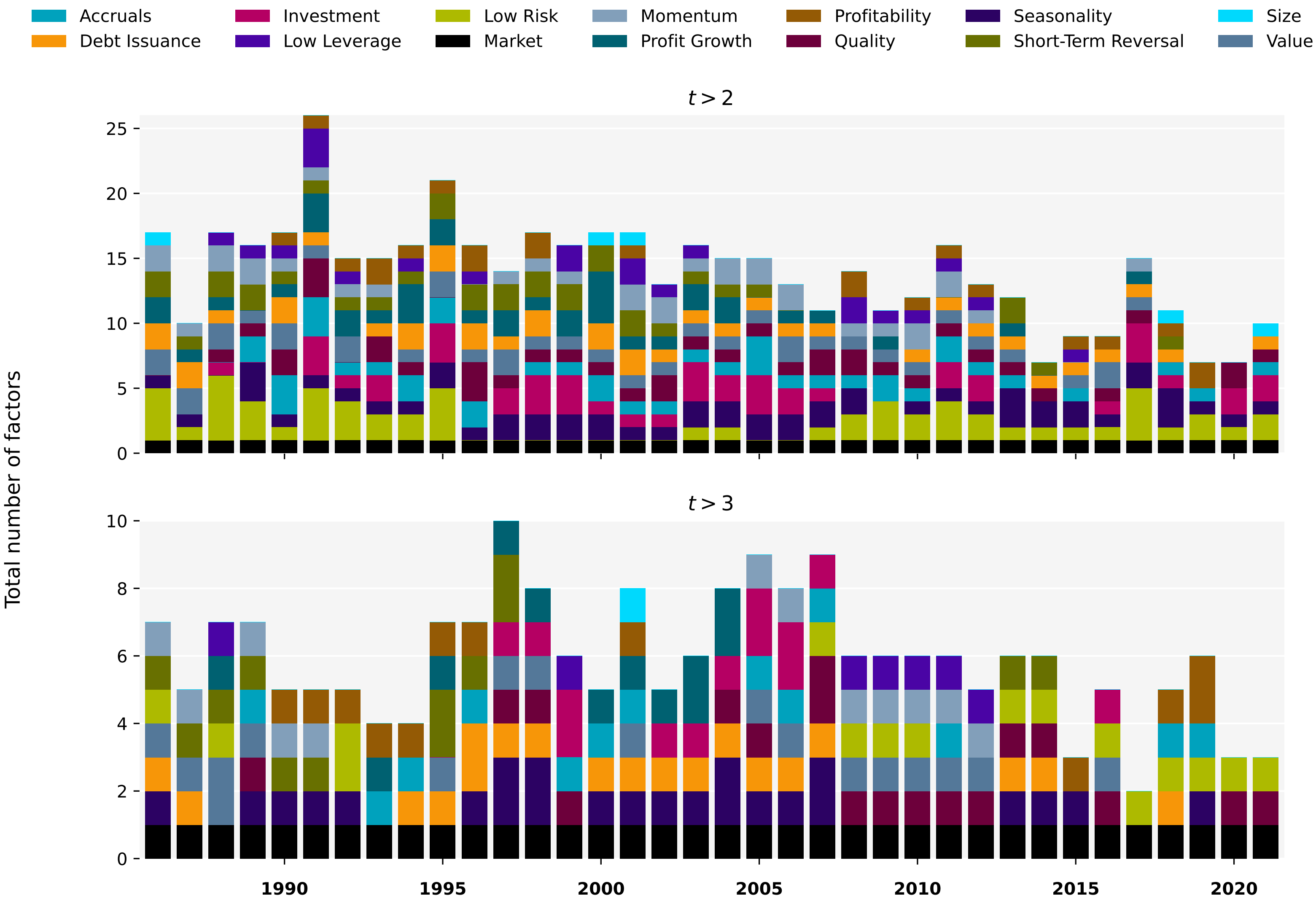
This figure depicts annualized CAPM-alphas for the U.S. factor zoo. The factors selected by the iterative factor selection process are indicated by full colors; all other factors are whitened. The underlying excess market return is sourced from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.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

Figure 3: Factor persistence



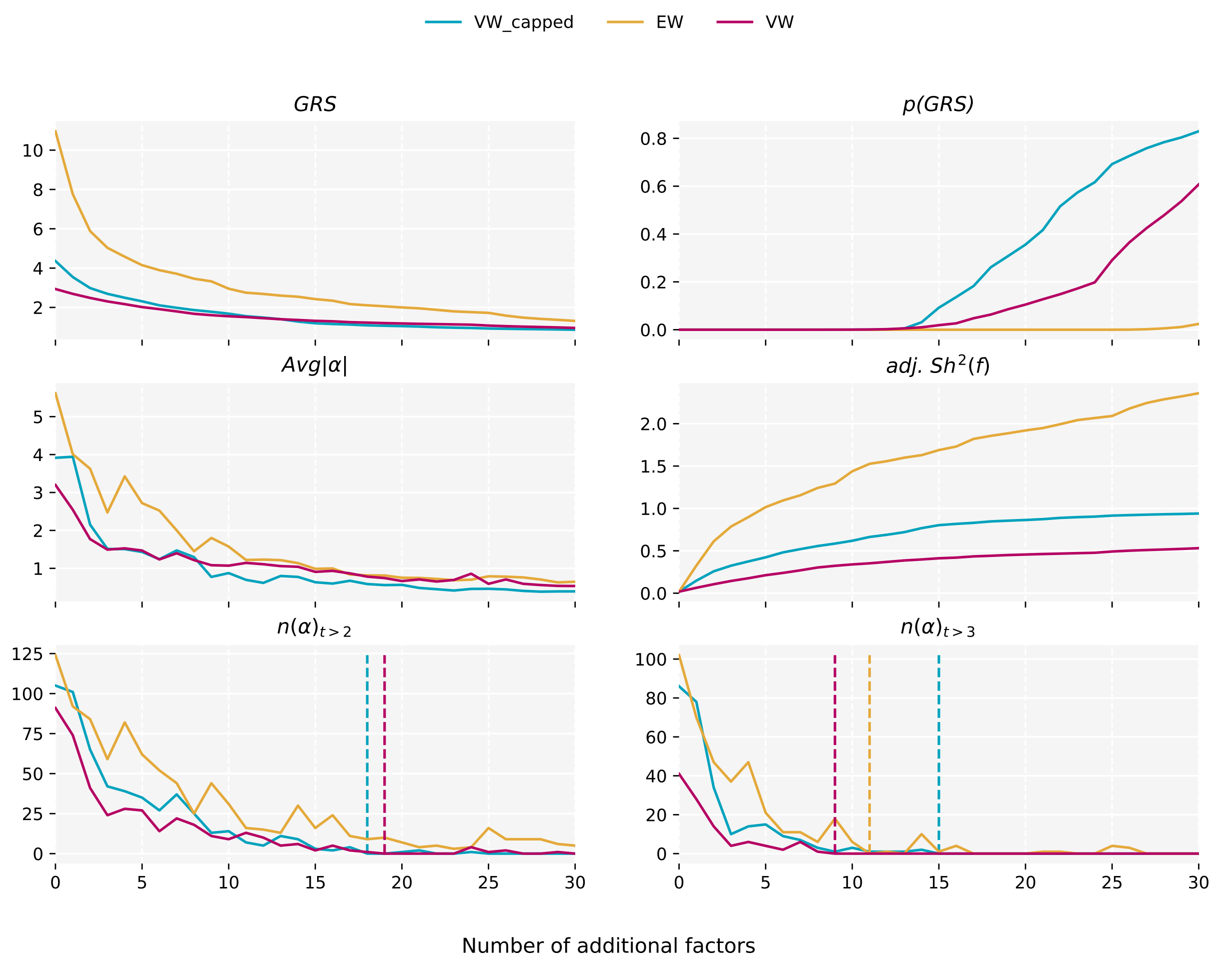
This figure depicts the evolution of the iterative factor selection process over time when utilizing a significance threshold of $t > 3.00$. Factors are added to the factor zoo on an annual basis following their publication. The publication year for each factor is indicated by a star. The initial estimation window size is 180 months and is expanded in subsequent iterations. The sample period is November 1971 to December 2021.

Figure 4: Rolling window factor selection



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

Figure 5: Alternative factor weighting schemes



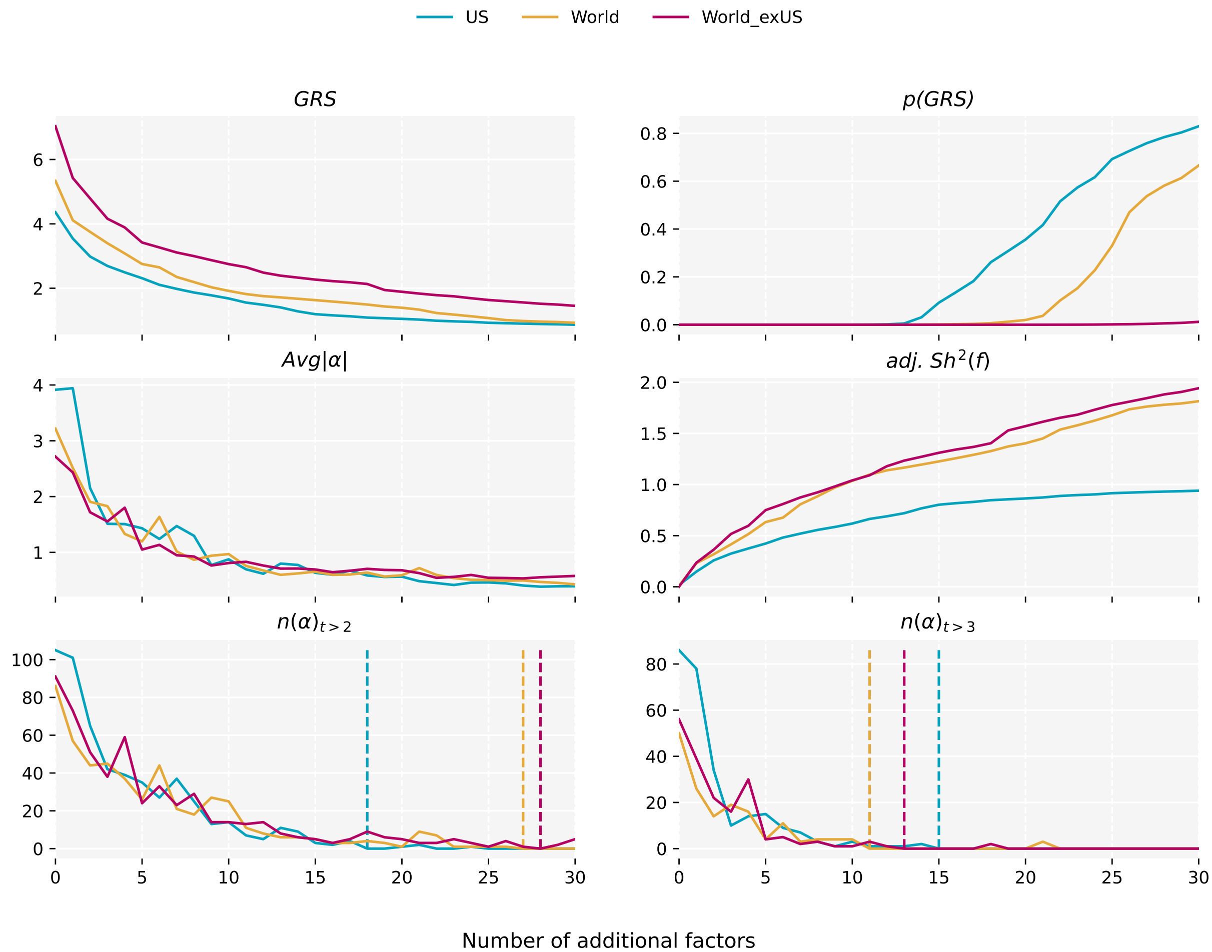
This figure depicts the key performance statistics for iterative factor models when based on different weighting schemes. We consider equal-weighting (EW), value-weighting (VW), and capped value-weighting (CW). The number of factors refers to the model building process described in Section 2. Average absolute alphas $Avg|\alpha|$ are annualised and in percentage. The vertical dashed lines in the last panel mark the minimum amount of factors needed to explain away the available factor zoo alpha in each weighting setting. The sample period is November 1971 to December 2021.

Table 3: Global factor analysis

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

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

Figure 6: Different regions



This figure depicts the key performance statistics for iterative factor models in different regions. The number and choice of factors refers to the model building process described in Section 2. Average absolute alphas $Avg|\alpha|$ are annualised and in percentage points. The vertical dashed lines in the lower panel mark the minimum amount of factors needed to explain the available factor zoo alpha. The sample period is August 1993 to December 2021 and considers 136 common factors for all three regions.

References

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

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

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

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

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

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

Bessembinder, Hendrik, Aaron Burt, and Christopher M. Hrdlicka (2022). "Factor returns and out-of-sample alphas: Factor construction matters." *Available at SSRN: https://ssrn.com/abstract=4281769*.

Blitz, David (2023). "The cross-section of factor returns." *Journal of Portfolio Management, forthcoming*.

Blitz, David, Matthias X. Hanauer, Iman Honarvar, Rob Huisman, and Pim van Vliet (2023). "Beyond Fama-French factors: Alpha from short-term signals." *Financial Analysts Journal* 79 (4), 96–117.

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

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

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

Chen, Andrew Y. and Tom Zimmermann (2022). "Open source cross-sectional asset pricing." *Critical Finance Review* 11 (2), 207–264.

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

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

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

Fama, Eugene F. and Kenneth R. French (1993). "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics* 33 (1), 3–56.

Fama, Eugene F. and Kenneth R. French (1996). "Multifactor explanations of asset pricing anomalies." *Journal of Finance* 51 (1), 55–84.

Fama, Eugene F. and Kenneth R. French (2012). "Size, value, and momentum in international stock returns." *Journal of Financial Economics* 105 (3), 457–472.

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

Fama, Eugene F. and Kenneth R. French (2016). "Dissecting anomalies with a five-factor model." *Review of Financial Studies* 29 (1), 69–103.

Fama, Eugene F. and Kenneth R. French (2018). "Choosing factors." *Journal of Financial Economics* 128 (2), 234–252.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken (1989). "A test of the efficiency of a given portfolio." *Econometrica*, 1121–1152.

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

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

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

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

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

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

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

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

Hou, Kewei, Chen Xue, and Lu Zhang (2020). "Replicating anomalies." *Review of Financial Studies* 33 (5), 2019–2133.

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

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

Jensen, Theis Ingerslev, Bryan T. Kelly, and Lasse Heje Pedersen (2023). "Is there a replication crisis in finance?." *Journal of Finance* 78 (5), 2465–2518.

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

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

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

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

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

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

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

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

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

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

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

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

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