Global Factor Premiums

Guido Baltussen*

Laurens Swinkels

Pim Van Vliet

Abstract

We examine 24 global factor premiums across the main asset classes via replication and new-sample evidence spanning more than 200 years of data. Replication yields ambiguous evidence within a unified testing framework with methods that account for p-hacking. The new-sample evidence reveals that the large majority of global factors are strongly present under conservative p-hacking perspectives, with limited out-of-sample decay of the premiums. Further, utilizing our deep sample, we find global factor premiums to be not driven by market, downside, or macroeconomic risks. These results reveal strong global factor premiums that present a challenge to asset pricing theories.

First version: April 2017

This Version: January 2019

Keywords: factor premium, multiple hypotheses testing, p-hacking, return anomalies, predictability, stocks, bonds, currencies, commodities, value, momentum, trend, carry, betting-against-beta, seasonality.

JEL classification: C11, C12, F31, G11, G12, G15, N210

*Baltussen (corresponding author) and Swinkels are from the Erasmus School of Economics, Erasmus University Rotterdam and Robeco Institutional Asset Management. Email: baltussen@ese.eur.nl Van Vliet is from Robeco Institutional Asset Management. We would like to thank David Blitz, Andrea Frazzini, Winfried Hallerbach, Matthias Hanauer, Matti Keloharju, Ralph Koijen, Martin Martens, Lasse Pedersen, and Jeroen van Zundert for valuable discussions and useful feedback. The views expressed in this study are not necessarily shared by Robeco Institutional Asset Management. We welcome comments, including references to related papers we have inadvertently overlooked.

I. Introduction

In this paper we study global factors premiums over a long and wide sample spanning the recent 217 years across equity index (but not single securities), bond, currency, and commodity markets. Several recent influential studies document the presence of significant factor premiums across these major asset classes. Moskowitz, Ooi, and Pedersen (2012) show the presence of a 'time-series momentum' premium, Asness, Moskowitz, and Pedersen (2013) reveal cross-sectional 'momentum' and 'value' premiums, Koijen, Moskowitz, Pedersen, and Vrugt (2018) find a 'carry' premium, Keloharju, Linnainmaa, and Nyberg (2016) discover a return seasonality premium, and Frazzini and Pedersen (2014) document a 'betting-against-beta' premium. The sample periods in these studies start typically around 1980 and confirm and extend earlier empirical asset pricing studies which often focus on a single asset class, usually U.S. equities.[1]

The first objective of this study is to robustly and rigorously examine these global factor premiums from the perspective of 'p-hacking' (see Harvey, 2017). As scientists, we are subject to statistical testing limitations, biased in the return anomalies we see published, have several degrees of freedom (on areas like data manipulation, statistical method, aggregation schemes, results we choose to present), and an incentive to publish. Consequently, several major findings might reflect (a combination of) cherry-picking the most significant results, a publication bias, a multiple hypothesis testing bias, and a type I error in testing (i.e. falsely discovering predictability). As Harvey (2017) notes: "*Given the competition for top journal space, there is an incentive to produce "significant" results. With the combination of unreported tests, lack of adjustment for multiple tests, and direct and indirect p-hacking, many of the results being published will fail to hold up in the*

[1] Several papers show the presence of return factors for individual asset classes. For equities, 'value', 'momentum', and 'low-risk' are well documented (e.g. Fama and French 1993, Jegadeesh and Titman, 1993, Blitz and Van Vliet, 2007). For currencies, Froot and Thaler (1990) and Barroso and Santa-Clara (2015) document a 'carry' factor, Menkhoff et al. (2012) document a 'momentum' factor, while Abuaf and Jorion (1990) and Menkhoff et al. (2017) document a 'value' factor. For commodities, Erb and Harvey (2006) document the 'momentum' factor. For bonds, Fama (1984) documents the term premium, also referred to as 'carry', and Ilmanen (1995) documents a 'value' and 'carry' factor.

future."[2] As a case in a point, Harvey, Liu, and Zhu (2016) find a clear publication bias pattern in the top finance journals, and that of over 300 documented stock-level anomalies many become questionable after analyzing these in a rigorous testing framework that allows for multiple hypotheses testing bias.[3]

We take as our starting point the main global return factors published in the Journal of Finance and the Journal of Financial Economics during the period 2012-2018: time-series momentum (henceforth 'trend'), cross-sectional momentum (henceforth 'momentum'), value, carry, return seasonality and betting-against-beta (henceforth 'BAB'). We examine these global factors in four major asset classes: equity indices, government bonds, commodities and currencies, hence resulting in a total of 24 global return factors.[4] We work from the idea that these published factor premiums could be influenced by p-hacking and that an extended sample period is useful for falsification or verification tests. Figure 1, Panel A summarizes the main results of these studies. Shown are the reported Sharpe ratio's in previous publications, as well as the 5% significance cutoff in the grey-colored dashed line. In general, the studies show evidence on the global factor premiums, with 14 of the 22 factors (return seasonality is not tested in bonds and currencies) displaying significant Sharpe ratio's at the conventional 5% significance level.

INSERT FIGURE 1 HERE

[2] P-hacking is not limited to financial economics. P-hacking is mostly discussed in social sciences and medicine. The Economist discussed the topic in 2013 with the headline title: '*How science goes wrong*'. Begley and Ellis (2012) show that out of 53 studies on pre-clinical cancer only 11% could be replicated. An open science collaboration (2015) shows that out of 97 significant psychological studies only 36 could be replicated. In behavioral economics, Camerer et al. (2016) find that out of 18 laboratory studies in economics 11 can be replicated.

[3] In a similar spirit, Hou, Xue, and Zhang (2017) find that 64% (85%) of almost 450 documented anomalies have t-statistics below two (three) when the importance of small and micro capitalization stocks is reduced. Chordia, Goyal, and Saretto (2017) show with a data mining approach that of about 2.1 million possible trading strategies only a small group survives after correcting for a multiple hypothesis testing bias, or as they state after using the "proper statistical hurdles". Moreover, the few surviving trading strategies seem to have no apparent theoretical underpinning. Further, Linnainmaa and Roberts (2018) show that when tested out of sample, many equity anomalies are weak and those that persist do so typically at about half of their original size.

[4] In this paper we focus on global market data in an uniform set of asset classes, and do not include individual U.S. stocks, international stocks or other company-level securities.

The question is what could be the impact of p-hacking on these numbers? To this end, we next turn to methods proposed against p-hacking. First, the black-colored dashed line in Figure 1 shows the 3.00 t-value cutoff advocated by Harvey, Liu, and Zhu (2016) and others.[5] Utilizing this cutoff only 10 of the 22 factor premiums are significant. Second, we apply the Bayesian perspective on p-values advocated by Harvey (2017), using a symmetric and descending minimum Bayes factor with "Perhaps" prior odds ratio 4:1. The numbers above each bar in Figure 1 indicate these "Bayesianized" p-values. Using this approach, we find that just 8 of the 22 factor premiums are significant.

Further, most of the studies have differences in, amongst others, testing methodologies, investment universes and sample periods, choices that introduce degrees of freedom to the researcher. To mitigate the impact of such degrees of freedom, we reexamine the global return factors using uniform choices on testing methodology and investment universe over their average sample period (1981-2011). Note that for uniformity this now includes testing return seasonality for bonds and currencies, which was not done in the original publication. Figure 1, Panel B shows the results of this replicating exercise. We find that Sharpe ratios are marginally lower, with 12 of the 24 factor premiums being significant at the conventional 5% level. Utilizing the 3.00 t-value cutoff, eight are significant, while only six factor premiums have Bayesian p-values below 5%.[6]

[5] Harvey, Liu, and Zhu (2016) conclude their analysis on the multiple testing bias in hypothesis testing in finance by advocating a t-statistic hurdle of 3 for factors to be deemed statistically interesting. Similarly, Benjamin et al. (2018) propose to redefine statistical significance across disciplines from the usual arbitrary p-value of 0.05 to an equally arbitrary, but stricter p-value of 0.005, thereby essentially increasing the t-statistic 'hurdle' from 1.96 to 2.81. They recommend employing statistical techniques to deal with multiple testing and Bayesian inference. However, this higher t-value is an easy step to implement, and the other techniques may be more technically demanding or there is no general agreement about their use. Harvey, Liu, and Zhu (2016)'s suggestion of increasing the t-statistic to 3 is thus very similar to the proposal of Benjamin et al. (2018).

[6] One could argue that another way to limit sensitivity to p-hacking is to combine the information of each factor across the different asset classes. An objection against such a multi-asset aggregation of evidence is that the factor premiums tend to be lowly correlated across asset classes. Nevertheless, when we apply such an aggregation, we find that four of these six multi-asset combinations exceed the 1.96 and 3.00 t-value cut-offs, while three of the six have Bayesian p-values below 5%.

One downside of the Bayesian approach above is the subjectivity required in the formulation of prior odds, subjectivity that typically has a substantial impact on the Bayesian p-value. To this end, we introduce a new perspective on p-hacking: the formulation of 'break-even' prior odds, or those prior odds at which the Bayesian p-value would equal the confidence level chosen. These break-even prior odds remove the need to specify the prior-odds (but do require the confidence level), and allow for an interpretation in terms of prior odds that would be required to just accept the alternative. Only in 3 out of 24 cases the break-even odds exceed 5. These break-even odds ratios imply that one does not need to be extremely skeptical to disregard the empirical evidence provided in the literature. This leaves us to conclude that further analysis to these global return factors is warranted using new and independent data, in line with the recommendation of amongst others Arnott, Harvey, and Markowitz (2018).

The second objective of this study is to provide rigorous new sample evidence on the global return factors. To this end, we construct a deep, largely uncovered historical global database on the global return factors in the four major asset classes. This data consists of pre-sample data spanning the period 1800- 1980, supplemented with post-sample data from 2012-2016, such that we have an extensive new sample to conduct further analyses. If the global return factors were unintentionally the result of p-hacking, we would expect them to disappear for this new sample period.

Our new sample findings reveal consistent and ubiquitous evidence for the large majority of global return factors. Figure 1, Panel C summarizes our main findings by depicting the historical Sharpe ratio's in the new sample period. In terms of economic significance, the Sharpe ratios are substantial, with an average of 0.41. Remarkably, in contrast to most out-of-sample studies (see for example Linnainmaa and Roberts, 2018), we see very limited 'out-of-sample' decay of factor premiums (the average in-sample Sharpe ratio is similar). In terms of statistical significance and p-hacking perspectives, 19 of the 24 t-values are above 3.0,19 Bayesian p-values are below 5%, and the break-even prior odds generally need to be above 9,999 to have less than 5% probability that the null hypothesis is true. Such extremely large odds imply that one needs to be extremely skeptical to disregard the empirical evidence provided in this study. As a main exception, the BAB effect is present in equity markets in both samples, but the evidence is less robust for the commodity and currency markets.[7] Return seasonality in government bonds and currencies are significant factor premiums and are an extension of the empirical asset pricing literature.[8]

Next, to mitigate the influence of p-hacking we report the robustness of the global return factors to common degrees of freedom to the researcher. To this end we employ rolling 10-year sample periods and robustness checks with regard to the various degrees of freedom in testing, such as the period of rebalancing, accounting for lagged implementation, the exact portfolio construction method, and the trimming of extreme positive returns. We find the global return factors to be robust to these choices.

The third objective of this study is to gain further insights into the economic explanations of the global factor premiums using our full sample period ranging from 1800 to 2016. We first test the uniqueness of the factor premiums, as true commonality might be observable especially over the longer run. Our findings reveal that most return factors are largely uncorrelated; nor do they span each other. Trend factors form an exception in that they encompass momentum factors, thereby extending the findings of Moskowitz, Ooi, and Pedersen (2012) to a deep sample. These findings suggest that it is difficult to formulate a single uniform explanation based on a common global component.

[7] This may not be that surprising, as the original study already reported weak evidence for these three asset classes. We also find BAB in the cross-section of country bonds to be below our statistical hurdles (as in Frazzini and Pedersen, 2014), but we do not include yield curves due to lack of deep historical data for different maturities. Frazzini and Pedersen (2014) find especially strong results for BAB applied to bonds across the curve within a single country which can explain differences in results. We therefore expect a stronger low-risk premium for bonds if these data were available.

[8] Five factors pass both the strict Bayesian p-value (8 out of 22) and the pre-sample test (18 out of 22): trend in equities, bonds and commodities, momentum in equities, and finally carry in equities and bonds.

Next, we test if the global return factors can be reconciled with market risk, downside risk and macroeconomic risks. The theory suggests that expected returns can vary due to market risk, downside risk (Bawa and Lindenberg, 1977) or macroeconomic risk (Chen, Roll, and Ross, 1986, Fama and French, 1989, Ferson and Harvey, 1991).[9] The common sample (1981-2011) based on the past decades might be biased, given that they were quite exceptional (no major wars, growing global prosperity and only a few large recessions or periods of social unrest), which severely limits the number of 'bad states'. Our extensive multi-century sample includes a substantial number of bad states, with for example many bear markets (43 years) or recessions (74 years). It therefore allows us to more deeply examine to which extent the global return factors can be explained by market risk, downside risk, or macroeconomic risk.

We find no supporting evidence for (unified) explanations of the global factor returns based on market, downside, or macroeconomic risk. Jensen's alphas give a similar picture to the Sharpe ratios. Moreover, downside risk has limited explanatory power for the global factor premiums and a downside risk pricing model is rejected by the data. For carry we find some sensitivity to downside risk, but this is at best only a partial and limited explanation, in line with the findings of Koijen et al. (2018). Finally, we find no clear evidence that macroeconomic risks drive the return on the global return factors when comparing macroeconomic states and unconditional and conditional macroeconomic risk tests as employed by Chordia and Shivakumar (2002), Griffin, Ji and Martin (2003) and Keloharju, Linnainmaa, and Nyberg (2016). First, we generally cannot reject the null hypothesis that the average returns on the global return factors are the same during global recessionary and expansionary periods or other proxies for good or bad macroeconomic states. Second, the global return factors bear basically no statistically or

[9] Asness, Moskowitz and Pedersen (2013) show that global macroeconomic variables are generally not related to value and momentum returns, while Keloharju, Linnainmaa, and Nyberg (2016) find that seasonality returns are hard to reconcile with macroeconomic risks. Koijen et al. (2018) report that carry return drawdowns are more likely to occur during global recession periods.

economically significant relation to common global macroeconomic factors. Third, the predicted payoffs on the global return factors due to macroeconomic variables (e.g., based on the conditional forecasting approach proposed by Chordia and Shivakumar, 2002) can at best explain a small fraction of the total payoffs.

This study is not the first to utilize deep historical samples to study market premiums. For equity and bond premiums, Siegel (1992) gives evidence stretching back to 1800, Goetzmann (1993) to 1695, and Golez and Koudijs (2018) go even further back to 1629. Hurst, Ooi, and Pederson (2017) find a persistent trend premium going back to 1880, while Goetzmann and Huang (2018) show that stock-level momentum worked in imperial Russia in the period 1865-1914. Further, Doskov and Swinkels (2015) and Taylor (2002) find that carry and value premiums are present in currency markets since 1900. These historical studies typically focus on a single factor (e.g. momentum), a single market (often the U.S.), or a single asset class (typically stocks) and typically employ different methodologies. In this study, we focus on a broad range of global return factors across a broad range of markets using an uniform testing methodology and addressing p-hacking concerns.

The remainder of this paper is structured as follows: Section II describes the original studies and the results of our replication exercises from an economic and p-hacking perspective. It also introduces the Bayesian concept of break-even prior odds. In Section III, we introduce our newly constructed deep historical database that starts in 1800, which allows us to evaluate global factor returns over a period that has mostly been unexplored in previous research. Section IV tests economic explanations of global return factors using this deep sample. Section V presents our conclusions. The appendices contain more details about the data and portfolio construction methods, as well as additional results.

II. Replicating and p-hacking insights on the recent sample

We start our study by summarizing the main results and testing choices made in the original studies, followed by our replication exercise. We incorporate economic, traditional statistical, and Bayesian perspectives that account for p-hacking. Further, we introduce a new perspective on p-hacking; the concept of break-even prior odds.

A. Original studies

We base our paper on five key studies that were published in the Journal of Finance and the Journal of Financial Economics during the period 2012-2018, documenting global return factors across the major asset classes: trend, momentum, value, carry, return seasonality, and BAB. Throughout this study, we choose to focus on broad assets within four major asset classes generally shared by these five key studies: international equity indices, 10-year government bond indices, commodities, and currencies.[10] Table I summarizes the main findings of, and testing choices made by, these studies for the four asset classes that we consider. The table further contains the definitions used for each factor within each asset class, the reported statistical and economic significance, the sample period, the portfolio construction method, and the number of assets within each asset class.

INSERT TABLE I HERE

As becomes clear from Table I, these studies report evidence supporting the global return factors across the four major asset classes we study. The Sharpe ratios for each of the 22 global return factors are positive (return seasonality is not tested in bonds and currencies), with many above 0.3, while t-values are generally well over two for 14 out of

[10] Note that we ignore asset classes that lack deeper historical data, but are included in one or two specific studies. These include, for example, credit markets and option markets. Further, we choose to focus on global markets, leaving the most widely studied asset class, single name stock markets, to further study.

22 factors. Especially carry and trend have high t-values for each individual asset class, corresponding to high conventional levels of significance. Value is relatively weaker, with t-values below 1.96 for three out of four asset classes.

In addition, although the studies have many similarities, several degrees of freedom in testing are accommodated differently (driven by sample selection choices, data limitations, etc.). For example, the start dates range from 1972 for equity value and cross-sectional momentum, to 1989 for BAB for bond indices and commodities. The portfolio construction method for trend strategies is required to be different from the other factors, but for the cross-sectional anomalies there are also two different methods: forming equally weighted portfolios for the top and bottom terciles and comparing the returns between these two portfolios, or taking a position proportional to the rank of the asset in the cross-section, with long positions for assets above the median, and short positions for those below the median. Moreover, the maximum number of assets varies across studies.

B. Replication

Next, we replicate these original studies in a unified testing framework in which we limit the degrees of freedom. We utilize (i) a uniform sample period, (ii) a uniform cross-section of assets, and (iii) a uniform factor construction method for the cross-sectional factors. However, we verify that the results below are robust and stand up to variation in the settings shown in Table I, as shown in Appendix Table B.1. We take as our sample period the average start and end years used in the original studies: 1981-2011.

More specifically, we consider four major global asset classes in developed markets: international equity indices, 10-year government bond indices, commodities, and currencies. Within each asset class we construct factor portfolios on each of the factors (the definitions and their motivations can be found in Appendix A.I). In a nutshell, trend and momentum are defined as the 12-month-minus-1-month excess return. Value is defined as dividend yield for equity indices, real yield for bonds, 5-year reversal in spot prices for commodities, and absolute and relative purchasing power parity for currencies. Carry is defined as the implied yield on each instrument. Return seasonality is defined as the return on an asset in a certain month over the prior 20 years. BAB is long the low beta assets and short the high beta assets with positions neutralized for the ex-ante beta relative to the global asset class portfolio return.

We construct factor investment portfolios at the end of every month in the spirit of their original papers. For the trend factor, which is directional in nature, we go long (short) markets in each asset class when the trend measure is positive (negative), following Moskowitz, Ooi, and Pedersen (2012). For the other factors, which are all cross-sectional in nature, we rank the markets in each investment universe based on the factor measure and take a position equal to the rank minus its cross-sectional average. This procedure is similar to that used by Asness, Moskowitz, and Pedersen (2013), Frazzini and Pedersen (2014), and Koijen et al. (2018). Further, we scale all positions and factors to a constant ex-ante volatility or by the ex-ante beta (the latter in the case of BAB). Our dataset is a collection from Bloomberg, Datastream and the OECD website. All returns we consider are in excess of local financing rates and expressed in U.S. dollars. Appendix A describes the data, factor construction, and sample choices in detail.

With these uniform factor definitions and portfolio construction rules, we replicate the 22 global factor premiums documented in the literature for the period 1981 to 2011. Note that for uniformity this now also includes testing return seasonality for bonds and currencies, which was not done in the original publication.

INSERT TABLE II HERE

The results of the replication are displayed in Table II. We find that the Sharpe ratios are marginally lower than in the original studies, averaging to 0.41, with 12 of the 24 factor premiums being significant at the conventional 5% level. The seasonal strategy applied to bond and currency markets, which was not reported in the original paper, is not significant during the period 1981 to 2011. Further, 8 of the global factor premiums are significant using a t-value of 3.00 as cut-off.

One could argue that another way to limit sensitivity to p-hacking is to combine information across asset classes. An objection against such a multi-asset aggregation of evidence is that the factor premiums tend to be lowly correlated across asset classes, as we will show in Section IV. Nevertheless, we also construct equal-volatility-weighted global factor portfolios per asset class and across asset classes ('Multi Asset') by targeting each market, then asset class portfolio, and then Multi Asset portfolio at a 10% ex-ante annual volatility and applying equal-weighting. Following this procedure, the Sharpe ratios of the six multi-asset combinations vary between 0.39 and 1.15, and the t-values are above two for each factor. The proposed threshold t-value of three is achieved by four out of six.

C. Statistically accounting for p-hacking

In this subsection, we report traditional p-values (which do not correct for multiple hypotheses testing bias), Bayesianized p-values advocated by Harvey (2017), and the 'break-even' prior odds. The critique on frequentist p-values is that they indicate how likely it is to observe the data under the assumption that the null hypothesis is true. However, we are typically more interested in knowing the probability that the null hypothesis is true given that we observed the data. This question can be answered by a Bayesian approach. However, a full-blown Bayesian analysis, in which we specify priors on all possible hypotheses and calculate posterior probabilities given the observed data for inference, is challenging. A sensible alternative is to use the Minimum Bayes Factor (MBF; Edwards, Lindman, and Savage, 1963). The Bayes Factor connects the prior odds (before having seen the data) with the posterior odds (after having seen the data). The MBF gives maximum advantage to the alternative hypothesis, and therefore represents the maximum amount the data can move the posterior odds away from the prior odds. The MBF is global in the sense that it accepts all possible alternative hypotheses as potential candidates. If we are prepared to assume that the prior probabilities of alternatives should be symmetrical and descending (SD) around the null hypothesis, which seems to be a reasonable assumption for many applications in finance, we obtain an SD-MBF (Bayarri and Berger 1998)[11], defined as:

$$\text{SD-MBF} = -\exp(1) \times \text{p-value} \times \ln(\text{p-value}) \tag{1}$$

Consequently, we can transform the frequentist p-value to a Bayesian p-value using a level of prior odds of the null being true relative to an alternative:

$$\text{Bayesian p-value} = \text{SD-MBF} \times \text{prior odds} \,/\, (1+\text{SD-MBF} \times \text{prior odds}) \tag{2}$$

For example, for a situation in which the null and alternative are equally likely (i.e. even prior odds), the frequentist p-value of 5% is transformed into a Bayesian p-value of 29%. When the alternative is a 'long shot' with prior odds 99-to-1, the frequentist p-value of 5% will be Bayesian to 98%. In other words, if we are skeptical about the alternative, more convincing data is required to change our minds than if the alternative is more easily conceivable.[12]

[11] The MBF is $\exp(-Z^2/2)$ for a normal approximated z-score Z. This means that for a frequentist p-value of 5% the MBF equals 0.15, while this is 0.41 for the SD-MBF. The latter, which we use throughout our paper, is clearly more conservative as it leads to higher Bayesian p-values.

[12] As also stressed by Harvey (2017), the Bayes factor we employ is in principle a first hurdle to filter effects that are highly unlikely to be true. As a next step one could argue that a more conservative Bayes factor is needed, or even an explicit alternative. However, Bayarri, Benjamin, Berger, and Sellke (2016) show that the SD-MBF is generally very close to the full Bayes factor when p-values are low (as in our case).

The difficulty in the Bayesian approach is the assumption on the level of prior odds. This choice is quite subjective, while it typically has a material impact on the posterior p-value. As an alternative, we can think in terms of the 'break-even' prior odds, or that level of prior odds at which the Bayesian p-value equals the chosen significance level α. When we rewrite the equation above, we obtain the break-even prior odds as a function of α and the frequentist p-value (represented by the SD-MBF):

$$\text{Break-even prior odds} = \alpha / ((1 - \alpha) * \text{SD-MBF}) \tag{3}$$

Suppose we found a frequentist p-value of 0.1%, and want the α to be 5%, this implies a break-even prior odds ratio of 2.8:1. If we are *a priori* more skeptical about the alternative than this, the frequentist p-value of 0.1% (t-statistic of 3.3) do not provide enough evidence to change our minds.

INSERT TABLE III HERE

Table III shows the results of applying these Bayesian concepts to our replication exercise. For each global return factor, we report the frequentist p-values and the Bayesian p-values with a prior odds ratio of 4:1; prior odds classified by Harvey (2017) as 'perhaps' (see also Figure 1 for the results of both the original and replication study). We find that only 25% (six) of the 24 factor premiums, and 50% (three) of their six multi-asset combinations have Bayesian p-values below 5%. Next, turning to the break-even odds, we find (obviously) the same number of global return factors to have prior odds above 4:1. Turning to more conservative prior odds, we find that only in three out of 24 cases the break-even odds exceed 5, while only two exceed 99 – the prior odds labelled as a 'long shot' by Harvey (2017). These results imply that one does not need to be very skeptical to disregard the empirical evidence. This leaves us to conclude that further analysis is warranted using new and independent data for these factor premiums, a challenge we pick up next.

III. Global return factors: evidence since 1800

Next, we examine the global return factors using a novel, deep and largely unexplored historical sample that was not used for their original discovery. Using freshly uncovered historical data for cross-validation is one of the protocols advocated by Arnott, Harvey, and Markowitz (2018) to avoid the dangers of data mining.

A. Historical database construction

We have compiled our data from several sources in order to obtain a reliable and historically extensive dataset. Our sample covers 217 years of data from 31 December 1799 through 31 December 2016. We obtain the most recent historical data on financial market prices and macroeconomic series from Bloomberg, Datastream and the OECD website, and splice these before inception with data from (in order of preference): Global Financial Data, the Dimson, Staunton and Marsh Database, the Jordà-Schularick-Taylor Macrohistory Database[13], and/or Jeremy Siegel's website.[14] Our dataset construction for each asset and measure is described in more detail in Appendix A. Table A.1 in the appendix summarizes the markets included in our sample, the start dates of their excess return series, and the start dates of the value and carry factor measures per market.

INSERT FIGURE 2 HERE

Figure 2 shows that our sample consists of 13 global financial markets in 1800: 2 equity markets (U.S., U.K.), 3 bond markets (U.S., U.K., France), 5 commodities (Wheat,

[13] http://macrohistory.net/. See Jordà, Schularick and Taylor (2017) for more information.
[14] http://www.jeremysiegel.com/.

Cotton, Cocoa, Copper, Silver), and 3 currencies (GBP/USD, FRF/USD, USD/USD). In 1822 the sample increases to 18, and then gradually further over time, with several markets entering between 1860 and 1870 to increase the sample size to 36. At the start of World War I in 1914 there are 50 global assets. In 1974, the number of assets has increased to 66, and in 1999 the sample decreases from its maximum of 68 to 63, because of the introduction of the euro currency.[15]

Due to problems with exceptionally high levels of data uncertainty, we next exclude hyperinflation episodes from our sample by excluding assets from countries with ex-ante levels of monthly inflation over 50 percent. This definition follows Cagan (1956) and is widely accepted. We only start investing 12-months after the hyperinflation period has ended, as this is real-time available information. This definition of hyperinflation affects Germany during the 1920-1926 period; France in 1920-1921; France, Italy, and Japan during the post-World War II years (1946-1949/1950 and 1943-1948 for Italy); and South Korea in 1950-1956. Furthermore, in order to filter for liquidity and data historically available at a non-monthly frequency, we add a 'zero return screen', as assets with lower liquidity are more likely to have zero-returns. Lesmond, Ogden, and Trzcinka (1999) in essence show that the number of periods with zero-returns is an efficient proxy for liquidity. We include an asset at each point in time when over the past 12 months at most one month displayed a zero spot price movement or missing price. This screen has limited impact on our results, as shown in the robustness section.

In this study, we choose to focus on the main markets in each asset class based on what we can assess today, thereby ignoring the smaller markets. Most of these smaller markets were generally of lesser importance for investors, and hence this choice prevents us from finding factor premiums that would have been of small importance economically.

[15] Note that the limited number of assets per asset class in roughly the first 50 years of our sample period makes it more difficult to detect the existence of global return factors. Even though the average returns need not necessarily be affected, the variation around the average is probably higher due to limited diversification benefits in the portfolios.

Nevertheless, some of these markets might have been more relevant historically, possibly creating a survivorship bias as described in Brown, Goetzmann, and Ross (1995). For example, Argentina is commonly believed to have been an economy of comparable size to large European markets at the beginning of the 1900s, while our data sources do not cover equity or bond markets in Argentina around the 1900s and we do not include this market in our sample. That said, Dimson, Marsh, and Staunton (2008) claim that several Latin American countries combined (including Argentina) constituted less than 1.5% of the equity market capitalization during the 1900s. Moreover, Goetzmann and Kim (2018) analyze the equity indices in Global Financial Data for survivorship bias in a setting that is arguably more exposed to survivorship bias (i.e. crashes and rebounds) and show that this is at best a minor concern. Further, our results are robust across ten-year subperiods (including the recent subperiods), which seems hard to reconcile with survivorship bias.

Even though we (and the data vendors) have paid close attention to data quality, it is hard to assess to what extent these worries exactly apply, especially for data further back in history. In general, we expect that poor data quality would be random and therefore increase noise in the data, biasing our results towards the null hypothesis that the various global return factors do not exist. In addition to this qualitative argument, we examine robustness towards trimming extremely positive returns, applying an additional monthly lag between signal measurement and factor strategy implementation, and report results for rolling 10-year periods. These checks should mitigate concerns about the quality of our data.

B. Global return factors since 1800

The results of each global return factor, in terms of Sharpe ratios, over the 'new' period (1800-1980 and 2012-2016) are displayed in Panel A of Table IV.

INSERT TABLE IV HERE

First and foremost, we find a strong presence of most global return factors. In terms of economic significance, we find an average Sharpe ratio of 0.41 and in terms of statistical significance we find that 19 of the 24 t-values are above 3.0. Remarkably, unlike the results of most out-of-sample studies (see for example Linnainmaa and Roberts, 2018), we generally find very limited out-of-sample decay of the global return factors. In fact, the average Sharpe ratio of our replication exercise is also 0.41.

Panel B of Table IV exploits the full power of our sample by studying the full sample results. In general, results are stronger as measured by t-values because of the additional three decades with positive (although not always statistically significant) returns. Over both the new sample and the full sample, time-series momentum, carry, and seasonality are generally the strongest. For each of the four asset classes, these three factors have Sharpe ratios above 0.20 and t-values well above three (except for commodity carry in the new sample for which we only have data since 1968, yet over this relatively short period until 1980 the t-value is still 2.52). Return seasonality in government bonds and currencies are strong factor premiums and are an extension of the empirical asset pricing literature. For momentum and value we find statistically and economically significant positive returns for three out of four asset classes. Momentum for commodities and value for currencies is weak with t-values in the new (full) sample of -0.99 (0.71) and 0.36 (1.09), respectively.[16] The results for BAB reveal that the effect is strongly present for equity indices, but not significant for other asset classes.[17]

[16] We observe that our new sample for commodity momentum is based on spot returns, while the literature typically examines commodity futures returns instead.

[17] That said, our results for the BAB factor are comparable for all asset classes to Frazzini and Pedersen (2014) over the replication sample period. Table II reveals that, over the 1980-2011 period, BAB has an annualized Sharpe ratio of 0.37 for equities, 0.36 for country bonds, 0.01 for commodities and 0.13 for currencies. This compares to 0.51, 0.14, 0.11, and 0.22, respectively, documented in Frazzini and Pedersen (2014). Furthermore, Frazzini and Pedersen (2014) find substantially stronger results for BAB applied to bonds across the curve within a single country. In a separate U.S. bond sample covering many maturities they

INSERT TABLE V HERE

Table V shows the results of applying the Bayesian concepts to the new sample (Panel A) and the full sample (Panel B). In Table III we report for each global return factor the frequentist p-values, the Bayesian p-values with a prior odds ratio of 4:1, and the break-even prior odds that lead to a posterior probability of 5% for the null hypothesis. Over the full sample period (i.e. utilizing the full statistical power of our data) we find that 18 of the 24 global return factors have Bayesian p-values below 5%. In addition, the prior odds ratio generally needs to be above 99-to-1. This is even more true for the multi-asset combinations, for which the prior odds ratios are all above 9,999 to 1, except for BAB, for which we only find positive empirical evidence in the equity indices. These results imply that one needs to be extremely skeptical in order not to reject the null hypothesis that global factor returns are zero.

At this point, the following important aspect should be noted: The main purpose of this paper is to provide more robust and rigorous long-term evidence of the historical presence of global return factors, utilizing their most simple or basic definitions as put forward in influential papers analyzing recent samples. In this light, this study does not examine smarter and possibly better definitions, nor aspects linked to (limits to) arbitrage and tradability (such as transaction costs, turnover, legal controls, capital mobility, etc.). Although there are indications that some of the global return factors were exploited historically (e.g. Jobst, 2009), and their trading was feasible at limited transaction costs (e.g. Koudijs, 2016), we leave these aspects for future research. Instead, we have documented the consistent, robust existence of most global return factors over 217 years of data, which includes more than 160 years of pre-sample data.

find a Sharpe ratio of 0.81. Since bond data of various maturities is generally lacking deep history we have left the BAB factor in country bond markets outside the scope of this study.

C. Robustness analyses

An additional manner to mitigate the influence of p-hacking is to examine the robustness of the global return factors to common degrees of freedom for the researcher (e.g., sample periods, factor construction and other methodological choices). We start our robustness analyses with subperiod results. To reduce the degrees of freedom (that might lead to picking convenient subperiods), we consider all 10-year rolling subperiods, and report the frequency of positive Sharpe ratios in Table VI. The global return factors with significant full sample Sharpe ratios display consistent premiums over time, with frequencies typically above 80%. Further, with the exception of BAB, the multi-asset combinations have a positive Sharpe ratio in at least 90% of the rolling 10-year periods between 1800 and 2016. This consistency of performance over time further strengthens our empirical evidence for the existence of most of the global return factors we examine.

INSERT TABLE VI HERE

INSERT TABLE VII HERE

Next, Table VII considers methodological variations relative to our 'baseline strategies' examined so far. We focus on the full sample results, but we have verified that results are similar for the new sample period (see Table B.1 in Appendix B for robustness in the replication sample). First, we remove the zero-return liquidity screen, as this is not commonly employed in the original studies underlying this paper. Overall, removing this screen tends to increase the Sharpe ratios, most notably for value in currencies, which now becomes economically strong (0.51) and significant. Second, instead of using the portfolio weights that depend on the cross-sectional ranking, we take equally-weighted long and short positions in the top and bottom tercile within each asset class (the portfolio construction choice employed by Keloharju, Linnainmaa, and Nyberg, 2016). Third, we do not scale the positions using past volatility or beta (in the case of BAB), but rather take simple equal notional positions in each asset. This means that more weight is given to the riskiest assets within an asset class. Fourth, we consider robustness to signals lagging by one month before investing. This lowers the Sharpe ratio for each factor except value, but the economic and statistical magnitude of the reduction is limited for each factor except seasonality. By construction, this factor is linked to a monthly seasonal effect, and delaying the signal by a month actually reduces the average return to indistinguishable from zero. Fifth, instead of a monthly rebalancing of the factor portfolios we consider a quarterly rebalancing frequency. Note that here we define seasonality over quarterly instead of monthly periods. Sixth, we test robustness with respect to outliers by trimming extreme returns (which might be especially difficult to harvest). Note that this will also provide more robustness to potential data errors. More specifically, we trim all strategy returns before any position sizing in a single market to a maximum of 50%. To summarize the information in the robustness checks we report the average Sharpe ratio and the fraction of settings that yield a significant Sharpe ratio in the final two columns of Table VII. Overall, the significant results reported so far are robust with regard to the above variations, with average Sharpe ratios close to the baseline results and at least five of the seven variations yielding significant Sharpe ratios.

IV. Economic explanations of the global return factors

We have documented robust evidence for the presence of global factor premiums using replication and out-of-sample testing and utilizing various methods that account for p-hacking. In this Section, we turn to insights into potential explanations of these global factor premiums. In line with, amongst others, Koijen et al.(2018) and Keloharju, Linnainmaa, and Nyberg (2016), we consider common variation, and explanations related to market risk, downside risk, and macroeconomic risks.[18] To this end, we exploit the power of our full sample period ranging from 1800 to 2016. The advantage of this long-term data series is that it is easier to identify explanations that relate to relatively infrequent observations such as downside or macroeconomic risks.

A. Market risk and common variation

The Sharpe ratios for the factors that we have shown in the previous section measure return per unit of risk, but do not account for exposures to traditional asset classes. If the factor returns can be explained by correlation to traditional global markets, they are redundant, even if they have an economically significant Sharpe ratio. In order to adjust for these market exposures, we start by displaying the Jensen's alphas for each of the factor-asset class combinations, where we adjust the factor return series for exposures to market risks, in Table 8, Panel A. The results from these alphas are very comparable to those based on Sharpe ratios. The alphas are typically about 10 times larger than the Sharpe ratio, which one would expect if there is no correlation with global markets, as the factor returns are scaled to an (ex-post) volatility of 10% per annum. These results indicate that none of the factor strategies are strongly correlated to global asset class risks.

We continue in a next step to show the average pairwise monthly return correlations. These can be found in Panel B of Table 8. The average correlation of each factor across asset classes is close to zero, ranging from -0.01 (value) to 0.05 (trend). A similar picture emerges for the average correlation for factors within an asset class, with values ranging from 0.02 (commodities) to 0.10 (currencies). Furthermore, most individual correlation coefficients between the multi-asset factor series and between each factor-asset class

[18] Another explanation offered for many of the global return factors is market or funding liquidity risk (see amongst others Asness, Moskowitz and Pedersen, 2013, and Koijen et al., 2018). Due to the limited availability of deep historical data on the measures used in these studies we do not examine these explanations in this paper.

series are also close to zero, as shown in Appendix Table B.2 (containing the entire correlation matrix). From this perspective, most of the 24 global return factors are unique drivers of return, that share little common variation. The main exceptions are time-series and cross-sectional momentum that generally correlate positively, and both momentum factors and value correlate negatively.

To show the added value of each of the factors, we regress its monthly time-series on the 23 other return series (where we impute missing values with the average in order to exploit the full sample). The intercept of this regression can be interpreted as an expansion of the mean-variance efficient frontier with respect to the other factors (see De Roon and Nijman, 2001), and is hence a test of the unique added value of each of the 24 global return factors. All series are scaled to an ex-post volatility of 10% for interpretation purposes. Table VIII, Panel C shows the (annualized) intercepts from the spanning regressions.

In line with the above results, most intercepts are significantly positive. The main exception is that the trend factors subsume the momentum factors. The intercepts on the cross-sectional momentum are negative or statistically indistinguishable from zero (cross-sectional momentum in equity indices is the exception at 10% significance), while the intercepts on the trend factors are all significantly positive. These results are similar to the empirical results presented by Moskowitz, Ooi, and Pedersen (2012) over the period 1985 to 2009. In (unreported) results we have verified the robustness of this finding over subperiods. This leaves us to conclude that trend encompasses momentum. Furthermore, seasonality in currencies and BAB in equities are spanned by the other global return factors, while BAB in bonds has a significantly positive intercept at the 10% significance level. Seasonality in other asset classes (including bonds) remains economically and statistically significantly positive.[19]

[19] These results suggest long positions to most of the global return factors in optimal asset allocations. To check this, we have evaluated optimal long-only mean-variance asset allocations to traditional asset classes

INSERT TABLE VIII HERE

B. Downside risk

A large and growing literature considers whether various return anomalies compensate investors for downside or crash risk. For example, Doskov and Swinkels (2015) show that the currency carry factor is exposed to crash risk, and Koijen et al. (2018) show that downside risk captures a part of the carry return in fixed income and commodity markets. In this subsection, we consider downside risk explanations via the Downside Risk CAPM (DR CAPM) of Bawa and Lindenberg (1977). When the DR CAPM holds, assets with higher downside betas should have higher expected returns. Another way to interpret downside risk is as a conditional risk factor based on falling markets, also referred to as downstate beta.

Recently, Lettau, Maggiori and Weber (2014) have used the DR CAPM model and find that the carry factor has higher downside beta compared to the regular beta. Degrees of freedom in testing downside beta are the thresholds at which downside returns start to count, and the choice of benchmark return. Typical choices are zero, or lower thresholds deeper into the left tail of the return distribution, with threshold and equity market as benchmark. Furthermore, a common challenge when estimating downside risk exposures and premiums is the general reduction in the number of observations in the left tail (Post and Van Vliet, 2006). Especially market crashes do not happen very often. For example, Lettau, Maggiori and Weber (2014) use a threshold of -1 standard deviation of the equity market return, which results in 55 monthly observations out of 435 in their 1974-2010 sample (87% of all observations are excluded). In our 217-year sample, we have many

and the multi-asset combinations of the global return factors. Our findings reveal high allocation weights to each of the Multi Asset factors except cross-sectional momentum and BAB. To save space, we do not report these results separately.

more such events, allowing us to rigorously examine the hypothesis that downside risk can explain the premiums on the global return factors. For example, we have 43 years of equity bear markets and 218 downside market states in the case of a threshold of -1 standard deviation of the equity market return. This relatively large number of observations also enables us to study downside risk even further into the left tail of the distribution.

INSERT TABLE IX HERE

INSERT FIGURE 3 HERE

We apply both the CAPM and the DR CAPM using the settings used in Lettau, Maggiori and Weber (2014). Table IX summarizes the results in terms of CAPM alphas and betas, as well as the DR CAPM alphas and betas. Table IX is sorted according to their downside equity market betas, and significant alpha numbers are displayed in bold. Figure 3 depicts the excess return versus beta (right plot) or downside beta (left plot) of each global return factor. Overall, the average downside beta is very similar to the regular beta, with differences not larger than 0.10. Focusing on individual global return factors, carry has the highest increase in downside beta, in line with the results of Koijen et al. (2018) and Lettau, Maggiori and Weber (2014). Still, downside risk explains at best a small fraction of carry profits, with for example a downside beta - regular beta difference of 0.06 for the carry multi-asset combination, reducing its alpha by a marginal 0.22 percent to 9.73. In Appendix B, Table B.3 we further show the sensitivity of the downside beta results to using a zero-return cutoff or using equity crash betas (i.e. with the threshold deeper in the left tail). Moreover, Baltussen, Post and Van Vliet (2012) show that the stock-level value premium is exposed to downside risk in both equity and bond markets. Consequently, we also consider conditioning on bond market movements in

Table B.3. Overall, these results are similar to the above, with downside risk explaining at best a part of the global factor returns, most notably carry. Next, we estimate the risk premiums attached to beta or downside beta using the Fama and MacBeth (1973) approach, and find that the price of beta as priced in the cross-section of global return factors is of the wrong sign (-0.06) and insignificant (t-value = -0.39), as also evident from the flat line in Figure 3. When beta is replaced by downside beta in the Fama and MacBeth regressions we find an insignificant cross-sectional risk premium of 0.05 percent (t-value = 0.26). Moreover, instead of estimating the risk premium, we can infer it from the data as the time-series average equity premium (3.05 percent). Using this risk premium, the downside beta required to bring the average global factor return (4.11 percent) to zero would be above 1.30. This is a very high downside beta not observed for any of the factors in the full sample (for example, the maximum tail beta show in Appendix B is below 0.3). Moreover, we also do not observe such large downside betas for any of the 24 global return factors when considering any 10-year sample period (unreported for the sake of brevity). Based on these long-run sample results we conclude that downside market risk does not materially explain the global factor premiums.

C. Bad states and macroeconomic risk

Risk-based explanations of asset pricing anomalies argue for time-variation in expected returns related to time variation in its risk or risk premium, aspects that can be expected to relate to macroeconomic or market conditions (Chen, Roll, and Ross, 1986, Fama and French, 1989, Ferson and Harvey, 1991).[20] A major advantage of our extended sample (resulting from over 200 years of data) is the presence of many observations within

[20] For example, Merton's (1973) Intertemporal Capital Asset Pricing Model stipulates that in a risk-averse economy any variable that affects the set of future investment opportunities or the level of consumption earns a risk premium. Macroeconomic variables should be high up on the list of candidates, since they impact the cash flows of many agents in the real economy simultaneously, typically impact the real investment opportunities available (for example via government stimulus) and covary with risk appetite in the markets. Moreover, papers by, amongst others, Fama and French (1989) and Ferson and Harvey (1991) argue that expected business conditions are fundamental drivers of time variation in expected risk premiums.

and across various economic and market regimes (e.g., recessions and crisis episodes). A sample based on the past decades might be biased, given that they were quite exceptional (no major wars, growing global prosperity and only a few large recessions or periods of social unrest), which severely limits the number of 'bad states'. Our extensive multi-century sample includes many bear markets (43 years) and recessions (74 years). It therefore allows us to examine and describe in greater depth to what extent the global return factors can be explained with macroeconomic risk.

Many approaches have been used in the literature to examine sensitivity to, and pricing of macroeconomic risks. Following Griffin, Ji and Martin (2003) and Keloharju, Linnainmaa, and Nyberg (2016), we comprehensively examine macroeconomic risk explanations of the global return factors by taking three approaches. First, we divide periods into 'good' and 'bad' states and examine the pricing of global return factors over each. Second, we examine macroeconomic risk exposures and pricing using an unconditional approach, in the spirit of Chen, Roll and Ross (1986). Third, we examine time variation in global factor returns related to macroeconomic variables using a dynamic pricing approach, in the spirit of Chordia and Shivakumar (2002).

To examine the pricing of global return factors over good and bad states, we construct indicators of each market state and compute contemporaneous average annual factor returns for each state, as well as the return difference between the states. The market states, or regimes, that we examine are constructed per calendar year and are:

- Recession versus expansionary periods, where we mark calendar years as recessionary when at least six months of a calendar year are in a recession, and as expansionary otherwise. Our sample has 74 recession and 143 expansion years.
- Global crisis versus non-crisis periods, where we mark calendar years as 'Crisis' ('Non-crisis)' when the Rogoff and Reinhart Banking Currency Default Inflation (BCDI) indicator is above (below) 50. Our sample has 52 (165) crisis (non-crisis) years.

- Equity bull versus bear market periods, since one could argue that the equity markets provide a forward-looking assessment of macroeconomic conditions. We mark calendar years as a bull (bear) market when the calendar year global equity return series were positive (negative). Our sample has 43 (174) bear (bull) market years.

INSERT TABLE X HERE

INSERT TABLE XI HERE

The results are summarized in Table X. If the global return factors are driven by macroeconomic risks one would expect their returns to be especially strong in bad states. Overall, most global return factors do not display stronger returns during bad states (i.e. recessions, crises and turbulence). The most notable exception is value in equities during recession periods or bear markets, which also shows insignificant returns during expansion (but not bull markets). But these findings do not persist in crisis periods. The last row of Table X examines the common effects of macroeconomic states on all global return factors jointly, using a panel regression (with index-fixed effects and date-corrected standard errors). The results confirm significant global factor returns across all states, and only reveal significant variation across states for equity bear versus bull markets.[21]

[21] In addition, we have tested the impact of investor sentiment, as in Keloharju, Linnainmaa, and Nyberg (2016). Many studies argue that the significant presence of sentiment-driven investors can cause persistent mispricing and that profits on various return anomalies should be higher following high periods of sentiment if they are a reflection of mispricing (Baker and Wurgler, 2006, Stambaugh, Yu, and Yuan, 2012). We explore such an explanation by reconstructing the text-based market-wide pessimism measure of Garcia (2013) on a daily basis and extend it to span the 1899–2014 period (before 1899 data coverage is sparse). Subsequently, we average the word measures over each month, and divide our sample high and low sentiment months. In the spirit of Baker and Wurgler (2006) and Keloharju, Linnainmaa, and Nyberg (2016), we define high sentiment months as those in which the value of the market-wide pessimism index for the previous month is below the index's median value for the sample period. The low sentiment months are those that follow above-median pessimism index values. Our findings reveal that global return factors are generally significantly weaker in low sentiment states, and stronger in high sentiment states (but significantly positive in both), in line with the results of amongst other Stambaugh, Yu, and Yuan (2012), who find higher returns on stock-level return anomalies after periods of high sentiment. More specifically, we find that 14 out of 24 coefficients indicate stronger returns after high sentiment months (5 out of 6 for the multi-asset combinations), an effect that is strong and significant in the panel regression.

Our second analysis examines exposures to, and unconditional pricing of, macroeconomic factors. To this end, we construct the most widely used Chen, Roll and Ross (1986) factors – log changes in industrial production (MP; as in Chen et al. led by 1 month), term spread (UTS), changes in expected inflation (DEI), and unexpected inflation (UI) – for our global sample using monthly data.[22] We regress the time series of each global return factor on these macroeconomic variables and obtain coefficients and intercepts. Our sample starts in February 1869 due to the availability of deep historical inflation data. Table XI summarizes the results. If global return factors are driven by macroeconomic risk, then they should exhibit significant sensitivity to the factors proposed by Chen, Roll and Ross (1986). Our findings reveal that the global macroeconomic variables are generally not significantly related to global factor returns or subject to the wrong sign, with a couple of noteworthy exceptions. Value tends to load negatively on IP and UTS; carry and seasonality tend to load positively on IP but negatively on UTS and in the case of seasonality, also negatively on DEI. BAB tends to load positively on UTS. Value in equities has insignificant exposure to all macroeconomic factors. Furthermore, all significant global return factors of Section III have intercepts that are highly significant and are of similar magnitude to the raw returns over this sample (reported in the column "Actual"). These results suggest that macroeconomic risks have very limited explanatory power for the global return factors.

Next, to examine risk premiums attached to each macroeconomic factor and to what extent they can explain the global return factors, we apply the Fama and MacBeth (1973) technique on a monthly frequency with our global return factors as test assets. We combine the premiums with the estimated loadings to decompose the returns on the global

[22] More specifically, we construct these measures for the U.S., U.K., Germany, France and Japan (more details on the data can be found in Appendix A), and average these for a global measure. We have verified that our results are robust if U.S.-specific measures are used instead. Further, Chen, Roll and Ross (1986) show that these macroeconomic variables and the default premium are priced risk factors using the Fama and MacBeth (1973) regression on the cross-section of U.S. size portfolios. Akin to Griffin, Ji and Martin (2003), we omit the default premium, as its historical data availability is limited, especially outside the United States.

return factors into predicted and unexplained components. If the Chen, Roll and Ross factors suffice for explaining global return factors, then the difference between the actual and predicted returns (or unexplained) should not be significantly different from zero. The empirical results show that several global return factors have significant expected macroeconomic premiums, but that the unexplained return is generally more substantial and highly significant. The main exception is seasonality, especially in equity indices, where the unexplained return turns significantly negative.[23] However, this result originates from negative expected loadings combined with negative estimated risk premiums; the risk premiums (unreported) are either insignificant (in the case of IP and UI) or of the wrong sign (negative; UTS, DEI). When imposing risk premiums to be positive (unreported) this result disappears.

Our third analysis examines time variation in global factor returns related to macroeconomic variables. To this end we run the above time series regressions using a rolling 60-month window (as in Chordia and Shivakumar, 2002, and Griffin, Ji and Martin, 2003). We combine the estimated loadings with a realization on the macroeconomic variables to obtain one-step-ahead forecasts of returns, and compare the average expected and actual returns. The results, reported in the last two columns of Table XI, reveal at best a limited ability of the macroeconomic variables to predict global return factors, with most predicted returns being of the wrong sign (like seasonality in equities) or not significantly different from zero. In summary, our tests reveal very limited evidence of a link between macroeconomic risk and global return factors.

[23] Keloharju, Linnainmaa and Nyberg (2018) show that return seasonality in stocks is balanced by seasonal reversals: The seasonalities and seasonal reversals add up to zero over the calendar year. Their evidence suggests that return seasonalities are likely due to temporary mispricing.

V. Conclusion

We examine global factor premiums from three dimensions. First, we replicate existing studies to global factor premiums with a uniform factor construction methodology and sample period (thus limiting the degrees of freedom for each individual factor). Our results are generally quite similar to the published results, but methods to account for p-hacking reveal that more analyses is needed to draw conclusions on the existence of the majority of factor premiums. We introduce the Bayesian concept of break-even odds, the prior odds at which the Bayesian p-value would equal the confidence level chosen, which show that on average relatively low prior odds are needed to question the replication sample evidence on the global factor premiums.

Second, we build a wide and deep sample on the global return factors, covering more than 200 years of data across the major international asset classes. The new sample evidence reveals that the large majority of global factor premiums are convincingly present from economic, statistical, and p-hacking perspectives. Sharpe ratios are on average around 0.4 and generally persistent over time and factor construction settings. Further, most t-statistics well exceed 3, and break-even prior odds are generally above 9,999-to-1. Remarkably, the global return factors generally have close to zero decay in performance between the replication sample and out-of-sample periods. Factor premiums that are weak in the original studies are also weak in the extended sample period.

Third, we examine economic explanations for the global factor premiums. In general, the factors do not span each other, except for trend crowding out momentum. The global factor premiums are at best marginally explained by downside risk explanations. Furthermore, they are consistently present across various macroeconomic states, nor can be explained by macroeconomic risk models. Consequently, our results seem hard to reconcile with explanations based on risk, although we are cautious on such an interpretation because risk exposures and especially risk premiums are not directly observable. Instead, we interpret our findings as providing no positive evidence of a relationship between return factors and risk, nor on a unified explanation for the global return factors. Consequently, explanations of the documented global return factors are an important topic for future research.

Our results have strong implications for research in asset pricing. The literature on asset pricing theory and return predictability has often evolved separately by asset class or factor. Most studies focus on a single asset class, market or factor at a time, ignoring the wide presence of a group of return factors across asset classes. Our findings reveal that a theory that aims to model variation in expected returns should consider multiple asset classes and global return factors simultaneously. Furthermore, the documented return factors are important controls for empirical studies into new asset pricing factors. All in all, our study shows a strong, robust and persistent presence of economically important global factor premiums.

References

Abuaf, N., & Jorion, P. (1990). Purchasing power parity in the long run. *Journal of Finance*, 45(1), 157-174.

Arnott, R., Harvey, C., and Markowitz, H. (2018) A Backtesting Protocol in the Era of Machine Learning, *Journal of Financial Data Science,* forthcoming.

Asness, C.S., Moskowitz, T.J. and Pedersen, L.H. (2013). Value and momentum everywhere, *Journal of Finance*, 68(3), 929-985.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance* 61, 1645–1680.

Baltussen, G., Z. Da, & Van Bekkum, S. (2018). Indexing and Stock Market Serial Dependence around the World. *Journal of Financial Economics*, forthcoming.

Baltussen, G., Post, G.T. & Van Vliet, P. (2012). Downside Risk Aversion, Fixed-Income Exposure, and the Value Premium Puzzle. *Journal of Banking and Finance*, 36(12), 3382-3398.

Barroso, P., & Santa-Clara, P. (2015). Beyond the carry trade: Optimal currency portfolios. *Journal of Financial and Quantitative Analysis*, 50(5), 1037-1056.

Bayarri, M. J., Benjamin, D., Berger, J., & Sellke, T. (2016). Rejection odds and rejection ratios: A proposal for statistical practice in testing hypotheses, *Journal of Mathematical Psychology* 72, 90-103.

Bayarri, M.J., & Berger, J.O. (1998). Quantifying surprise in the data and model verification, in J.M. Bernardo, J.O. Berger, A.P. David, and A.F.M. Smith, eds: *Bayesian Statistics* 6, Oxford University Press: Oxford, U.K.

Bawa, V. S., & Lindenberg, E. B. (1977). Capital market equilibrium in a mean-lower partial moment framework. *Journal of Financial Economics*, *5*(2), 189-200.

Begley, C. G., & Ellis, L. M. (2012). Raise standards for preclinical cancer research. *Nature* 483, 531–533.

Benjamin, D., Berger, J., Johannesson, M., Nosek, B., Wagenmakers, E., Berk, R., Bollen, K., Brembs, B., Brown, L., Camerer, C., Cesarini, D., Chambers, C., Clyde, M., Cook, T., De Boeck, P., Dienes, Z., Dreber, A., Easwaran, K., Efferson, C., Fehr, E., Fidler, F., Field, A., Forster, M., George, E., Gonzalez, R., Goodman, S., Green, E., Green, D., Greenwald, A., Hadfield, J., Hedges, L., Held, L., Ho, T., Hoijtink, H., Hruschka, D., Imai, K., Imbens, G., Ioannidis, J., Jeon, M., Jones, J., Kirchler, M., Laibson, D., List, J., Little, R., Lupia, A., Machery, E., Maxwell, S., McCarthy, M., Moore, D., Morgan, S., Munafó, M., Nakagawa, S., Nyhan, B., Parker, T., Pericchi, L., Perugini, M., Rouder, J., Rousseau, J., Savalei, V., Schönbrodt, F., Sellke, T., Sinclair, B., Tingley, D., Van Zandt, T., Vazire, S., Watts, D., Winship, C., Wolpert, R., Xie, Y., Young, C., Zinman, J., & Johnson, V. (2018). Redefine statistical significance. *Nature Human Behaviour* 2, 6-10.

Blitz, D. C., & van Vliet, P. (2007). The Volatility Effect. *Journal of Portfolio Management*, 34(1), 102-113.

Brown, S., Goetzmann, W., & Ross, S. (1995). Survival. *Journal of Finance* 50(3), 853-873.

Cagan, P. 1956. The Monetary Dynamics of Hyperinflation. In *Studies in the Quantity Theory of Money*, ed. Milton Friedman, 25-117. Chicago: University of Chicago Press.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T. H., Huber, J., Johannesson, M., ... & Heikensten, E. (2016). Evaluating replicability of laboratory experiments in economics. *Science* 351(6280), 1433-1436.

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

Chordia, T., Goyal, A. & Saretto, A. (2017). p-Hacking: Evidence from Two Million Trading Strategies. SSRN working paper 3017677.

Chordia, T., & Shivakumar, L. (2002). Momentum, business cycle, and time-varying expected returns, *Journal of Finance* 57(4), 985-1019.

De Roon, F. A., & Nijman, T. E. (2001). Testing for mean-variance spanning: A survey. *Journal of Empirical Finance*, 8, 111-155.

Dimson, E., Marsh, P., & Staunton, M. (2008). The worldwide equity premium: A smaller puzzle. Chapter 11 in Mehra, R. (ed), *Handbook of the Equity Risk Premium*: Elsevier.

Doskov, N., & Swinkels, L. (2015). Empirical evidence on the currency carry trade, 1900–2012. *Journal of International Money and Finance*, 51, 370-389.

Edwards, W., Lindman, H., and Savage, L.J. (1963). Bayesian statistical inference for psychological research, *Psychological Review* 70, 193-242.

Erb, C. B., & Harvey, C. R. (2006). The strategic and tactical value of commodity futures. *Financial Analysts Journal*, 62(2), 69-97.

Fama, E. F. (1984). Term premiums in bond returns. *Journal of Financial Economics*, 13(4), 529-546.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25, 23-49.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56.

Fama, E. F., & Gibbons, M. R. (1984). A comparison of inflation forecasts. *Journal of Monetary Economics*, 13, 327-348.

Fama, E. F.; MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy* 81(3), 607–636.

Ferson, W. E. & Harvey, C. R., (1991). The Variation of Economic Risk Premiums. *Journal of Political Economy* 99, 385-415.

Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, 111(1), 1-25.

Froot, K. A., & Thaler, R. H. (1990). Anomalies: foreign exchange. *Journal of Economic Perspectives*, 4(3), 179-192.

Garcıa, D. (2013). Sentiment during recessions. *Journal of Finance* 68(3), 1267–1299.

Goetzmann, W.N., (1993). Patterns in three centuries of stock market prices. *Journal of Business*, 249-270.

Goetzmann, W.N., & Huang, S. (2018). Momentum in Imperial Russia, *Journal of Financial Economics* 130(3), 579-591.

Goetzmann, W.N., & D. Kim. (2018). Negative Bubbles: What Happens After a Crash. *European Financial Management* 24(2), 171-191.

Golez, B., & Koudijs, P. (2018). Four centuries of return predictability. *Journal of Financial Economics* 127(2), 248-263.

Griffin, J. M., Ji, X., & Martin, J. S. (2003). Momentum investing and business cycle risk: Evidence from pole to pole, *Journal of Finance* 58(6), 2515-2547.

Harvey, C.R., (2017). Presidential Address: The Scientific Outlook in Financial Economics. *Journal of Finance*, 72(4), pp. 1399–1440.

Harvey, C.R., Liu, Y. & Zhu, H. (2016). … and the Cross-Section of Expected Returns, *Review of Financial Studies*, 29(1), 5-68.

Hou, K., Xue, C., and Zhang, L. (2017). Replicating anomalies. *Review of Financial Studies,* forthcoming.

Hurst, B., Ooi, Y. H., & Pedersen, L. H. (2017). A century of evidence on trend-following investing. *Journal of Portfolio Management*, 44(1), 15-29.

Ilmanen, A. (1995). Time- varying expected returns in international bond markets. *The Journal of Finance*, *50*(2), 481-506.

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance* 48(1), 65-91.

Jensen, M.C. (1968). The Performance of Mutual Funds in the Period 1945-1964, *Journal of Finance* 23, 389-416.

Jobst, C. (2009). Market leader: The Austro-Hungarian Bank and the making of foreign exchange intervention, 1896-1913. *European Review of Economic History* 13(3), 287-318.

Jordà O., M. Schularick, & A. M. Taylor. 2017. Macrofinancial History and the New Business Cycle Facts. *NBER Macroeconomics Annual* 2016, volume 31, edited by Martin Eichenbaum and Jonathan A. Parker. Chicago: University of Chicago Press.

Keloharju, M., Linnainmaa, J. T., & Nyberg, P. (2016). Return seasonalities. *Journal of Finance*, 71(4), 1557-1590.

Keloharju, M., Linnainmaa, J. T., & Nyberg, P. (2018). Seasonal reversals in expected stock returns. *Working Paper*. Available at SSRN.

Koijen, R. S., Moskowitz, T. J., Pedersen, L. H., & Vrugt, E. B. (2018). Carry, *Journal of Financial Economics* 127(2), 197-225.

Koudijs, P. (2016). The boats that did not sail: Asset price volatility in a natural experiment. *Journal of Finance*, 71(3), 1185-1226.

Lesmond, D. A., Ogden, J. P. & Trzcinka, C. A. (1999). A new estimate of transaction costs, *Review of Financial Studies* 12, 1113-1141

Lettau, M., Maggiori, M., & Weber, M. (2014). Conditional risk premiums in currency markets and other asset classes. *Journal of Financial Economics*, 114(2), 197-225.

Linnainmaa, J. T., & Roberts, M. R. (2018). The History of the Cross Section of Stock Returns. *Review of Financial Studies* 31(7), 2606-2649.

Menkhoff, L., Sarno, L., Schmeling, M., & Schrimpf, A. (2012). Currency momentum strategies. *Journal of Financial Economics*, 106(3), 660-684.

Menkhoff, L., Sarno, L., Schmeling, M., & Schrimpf, A. (2017). Currency value. *Review of Financial Studies* 30(2), 416-441.

Merton, R. C., (1973). An intertemporal capital asset pricing model. *Econometrica* 41, 867-887.

Moskowitz, T.J., Ooi Y.H., & Pedersen, L.H. (2012). Time series momentum, *Journal of Financial Economics*, 104(2), 228-250.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716.

Post, T., & Van Vliet, P. (2006). Downside risk and asset pricing. *Journal of Banking & Finance*, 30(3), 823-849.

Siegel, J. J. (1992). The equity premium: Stock and bond returns since 1802. *Financial Analysts Journal*, 48(1), 28-38.

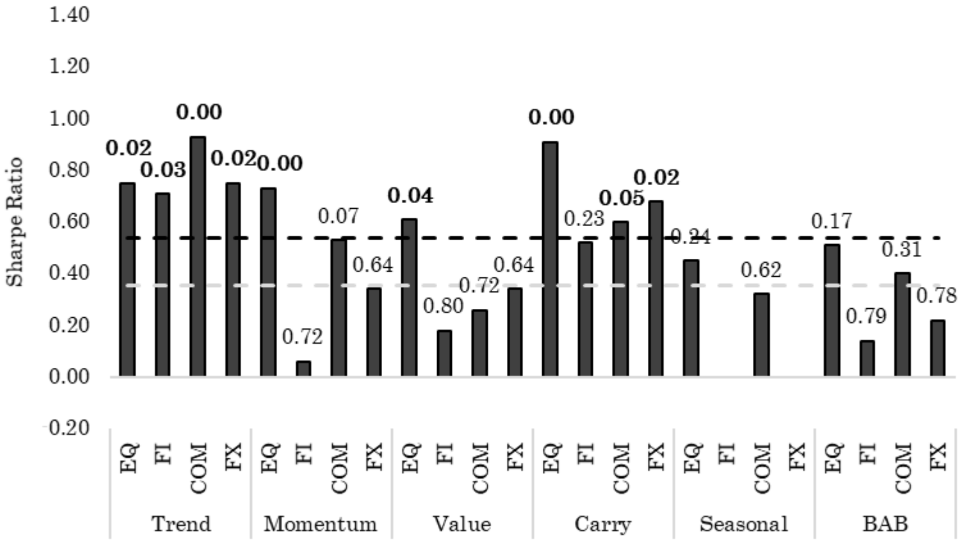
Stambaugh, R. F., Yu, J., & Yuan,Y., (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics* 104, 288–320.

Taylor, A. M. (2002). A century of purchasing-power parity. *Review of Economics and Statistics*, 84(1), 139-150.

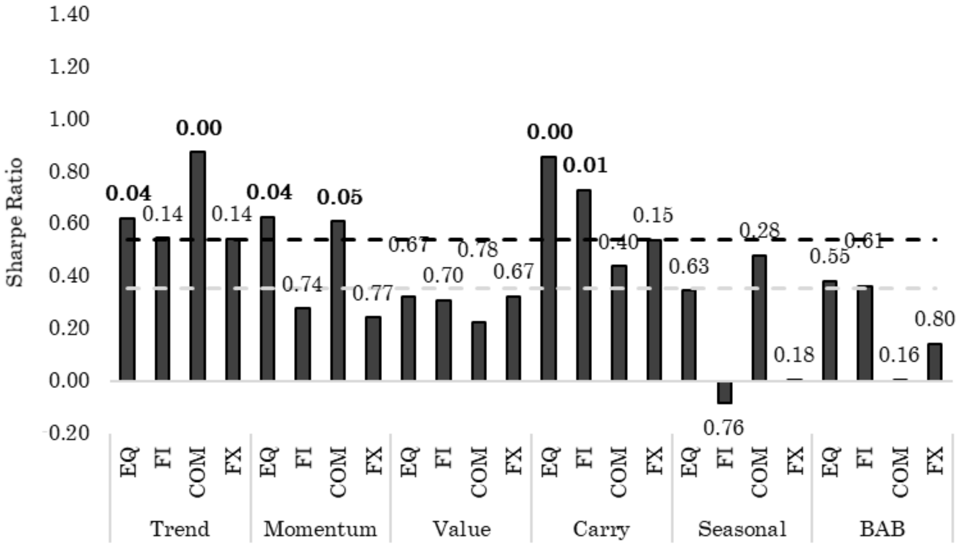
Figures

Figure 1: Global factor returns since 1800. The figure shows the annualized Sharpe ratios for the 24 global return factors. The dashed lines show the cutoff on the Sharpe ratio corresponding to t-values of 1.96 (in light grey) and 3.00 (in black) respectively. The values above each bar are the Bayesian p-values using a 4-to-1 prior odds ratio. Numbers in bold are below the 5% significance level. Panel A shows the results for the 'original sample' period, Panel B shows the results for the 'replication sample' period 1981-2011, Panel C shows the 'new sample' spanning the 1800-1980 and 2012-2016 period. The global factors are long-short portfolios applied on international equity indices ("EQ"), 10-year maturity government bond indices ("FI"), commodities ("COM"), and currencies ("FX").

Panel A: Original documentation



Panel B: Replicating factors 1981-2011



Panel C: New sample evidence: 1800-1980 & 2012-2016

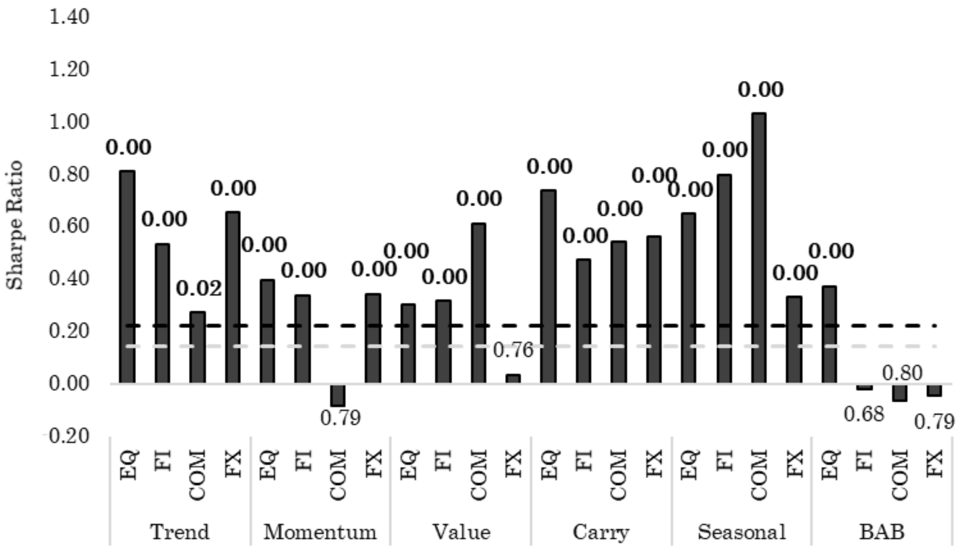


Figure 2: Global financial markets since 1800. The figure shows the number of markets included in the investment universe. The sample starts in December 1799 and ends December 2016 and is at the monthly frequency. Covered are equity indices ("Equities"), 10-year maturity government bond indices ("Bonds"), commodities ("Commodities"), and currencies ("FX").

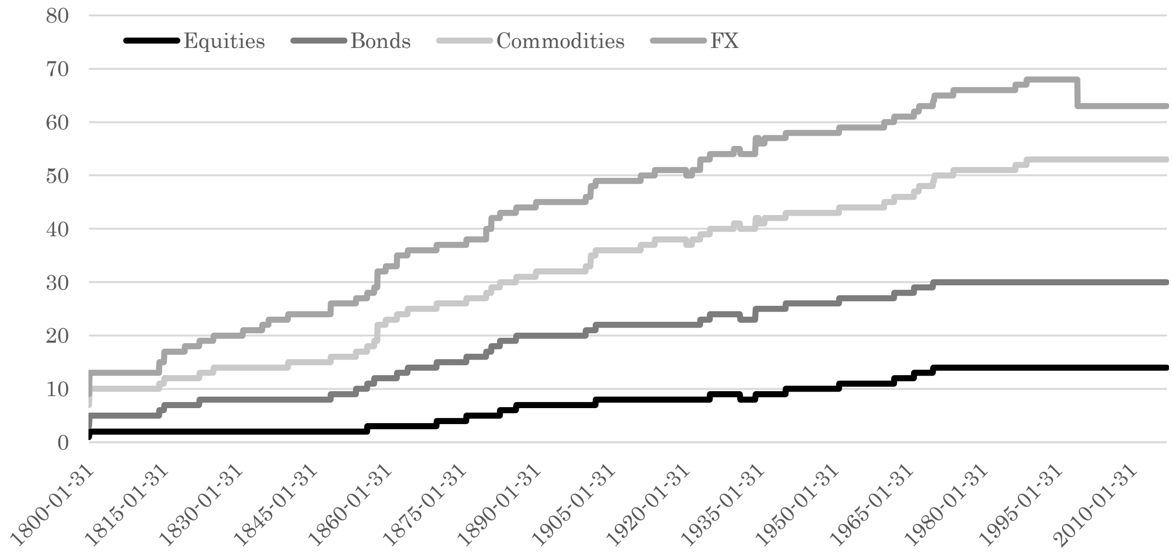
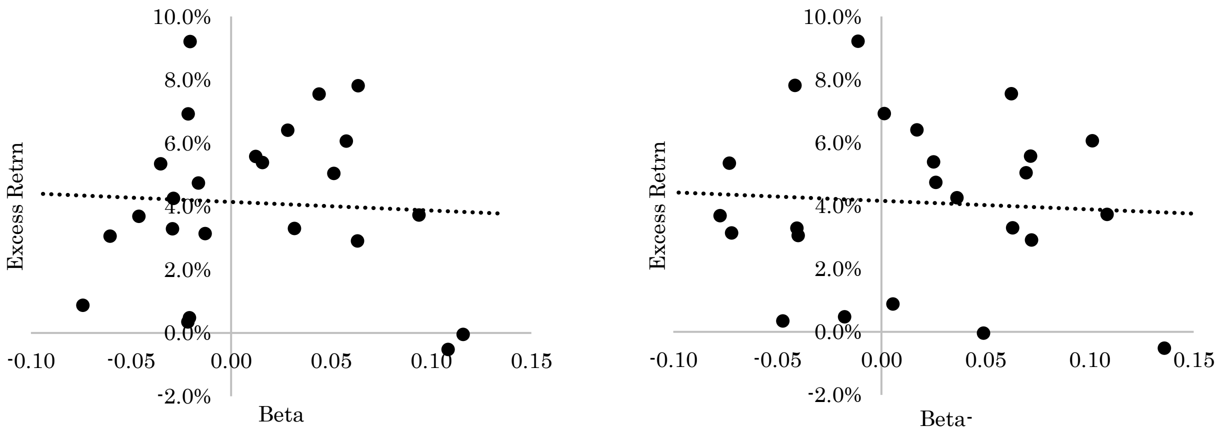


Figure 3: Downside risk and global factor premiums. The figure shows regular and downside betas versus the average return on the global factors. The left panel shows the regular beta ("Beta") and the right panel shows the downside beta ("Beta-"). The sample runs from 1800 until 2016 and is at the monthly frequency.



Tables

Table I: Original studies on global factor premiums

The table contains the factor definitions, t-stats, Sharpe ratios, start date, end date, portfolio construction method, and number of assets for each of the original studies on global factor premiums: Asness, Moskowitz, and Pedersen (2013) for value and cross-sectional momentum, Koijen, Moskowitz, Pedersen, and Vrugt (2018) for carry, Moskowitz, Ooi, and Pedersen (2012) for time-series momentum, Frazzini and Pedersen (2014) for the (betting against) beta anomaly, and Keloharju, Linnainmaa, and Nyberg (2016) for the monthly return seasonal. Abbreviations for the factor definitions are BE/ME: Book value of Equity divided by the Market value of Equity, 5Y Rel PPP: 5-year Relative Purchasing Power Parity, 5Y Rev: 5-year Reversal, 12M-1M: Past 12-month returns minus the last month, Impl DY: Futures-Implied Dividend Yield, Rate diff: Interest rate differential, 20Y,M: Average monthly return over the past 20-year period. Abbreviations for the portfolio construction are P1-P3: Equally-weighted top tercile portfolio minus the bottom tercile portfolio, Rank: Using weights proportional to the asset's rank, < >0: Volatility-weighted portfolio position depending on the sign of the past 12-month return. The number of assets is the maximum number of assets used per asset class. The columns with t-statistics or Sharpe ratios contain those reported as the main result; this does not exclude the possibility that results from other choices have also been presented in robustness analyses. The t-statistics or Sharpe ratios that are in italics are not presented in the original studies, but calculated using other information provided. Note that Asness, Moskowitz, and Pedersen (2013) report both P1-P3 and Rank weighted returns for value and cross-sectional momentum in their main table, and results are similar for the two methods.

Factor	Asset	Definition	t-stat	Sharpe ratio	Start	End	Portfolio	Assets
Trend	Equity indices	12M	3.77	*0.75*	1985	2009	<> 0	9
	Bond indices	12M	3.53	*0.71*	1985	2009	<> 0	13
	Commodities	12M	4.66	*0.93*	1985	2009	<> 0	24
	FX	12M	3.41	*0.68*	1985	2009	<> 0	12
Momentum	Equity indices	12M-1M	4.14	0.73	1972	2011	P1-P3	18
	Bond indices	12M-1M	0.35	0.06	1982	2011	P1-P3	10
	Commodities	12M-1M	3.29	0.53	1972	2011	P1-P3	27
	FX	12M-1M	1.90	0.34	1979	2011	P1-P3	10
Value	Equity indices	BE/ME	3.45	0.61	1972	2011	P1-P3	18
	Bond indices	5Y Rev.	0.97	0.18	1982	2011	P1-P3	10
	Commodities	5Y Rev.	1.61	0.26	1972	2011	P1-P3	27
	FX	5Y Rel. PPP	1.89	0.34	1979	2011	P1-P3	10
Carry	Equity indices	Impl. DY-rf	*4.46*	0.91	1988	2012	Rank	13
	Bond indices	Slope+Roll	*2.80*	0.52	1983	2012	Rank	10
	Commodities	Basis	*3.39*	0.60	1980	2012	Rank	24
	FX	Rate diff.	*3.66*	0.68	1983	2012	Rank	20
Seasonal	Equity indices	20Y, M	2.76	*0.45*	1974	2011	P1-P3	15
	Bond indices	20Y, M	-	-	-	-	-	-
	Commodities	20Y, M	1.97	*0.32*	1974	2011	P1-P3	24
	FX	20Y, M	-	-	-	-	-	-
BAB	Equity indices	Beta	2.93	0.51	1984	2012	Rank	13
	Bond indices	Beta	0.67	0.14	1989	2012	Rank	9
	Commodities	Beta	0.72	0.11	1989	2012	Rank	25
	FX	Beta	1.23	0.22	1977	2012	Rank	10
Average			**2.70**	**0.49**	**1981**	**2011**		**16**

Table II: Historical performance of global return factors: replicating sample

The table summarizes the historical performance of the global return factors. Shown per factor per asset class are the historical annualized Sharpe ratio. The sample starts in January 1981 and ends December 2011 and is at the monthly frequency. Covered are equity indices ("Equities"), 10-year maturity government bond indices ("Bonds"), commodities ("Commodities"), currencies ("FX"), and their equally weighted combination across the four asset classes ("Multi Asset"). Numbers in parentheses indicate t-values. Asterisks are used to indicate significance at a 10% (*), 5% (**) or 1% (***) level.

Sharpe ratio	Trend	Momentum	Value	Carry	Seasonality	BAB
Equities	0.62***	0.61***	0.34*	0.87***	0.36**	0.37**
	(3.44)	(3.39)	(1.89)	(4.83)	(2.00)	(2.05)
Bonds	0.53***	0.29	0.29	0.71***	-0.09	0.36**
	(2.94)	(1.61)	(1.61)	(3.94)	(-0.50)	(2.00)
Commodities	0.88***	0.62***	0.22	0.44**	0.48***	0.01
	(4.89)	(3.44)	(1.22)	(2.44)	(2.67)	(0.06)
FX	0.53***	0.23	0.32*	0.53***	0.00	0.13
	(2.94)	(1.28)	(1.78)	(2.94)	(0.00)	(0.72)
Multi Asset	1.09***	0.78***	0.55***	1.15***	0.40**	0.39**
	(6.05)	(4.33)	(3.05)	(6.39)	(2.22)	(2.17)

Table III: Statistical perspectives on global return factors: replicating sample

The table summarizes various statistical perspectives on the historical performance of the global return factors. Shown per factor per asset class are the historical frequentist p-value ("p-value"), Bayesianized p-value using a 4-to-1 prior odds ratio ("Bayesian-p") and break-even prior odds at a 5% confidence level ("BE-odds") of its performance. The sample starts in January 1981 and ends December 2011 and is at the monthly frequency. Covered are equity indices ("Equities"), 10-year maturity government bond indices ("Bonds"), commodities ("Commodities"), currencies ("FX"), and their equally weighted combination across the four asset classes ("Multi Asset").

		Trend	Momentum	Value	Carry	Seasonality	BAB
Equities	p-value	0.00	0.00	0.06	0.00	0.05	0.04
	Bayesian-p	0.04	0.05	0.64	0.00	0.60	0.58
	BE-odds	4.03	3.44	0.10	947.75	0.12	0.13
Bonds	p-value	0.00	0.11	0.11	0.00	0.62	0.05
	Bayesian-p	0.17	0.72	0.72	0.01	0.76	0.60
	BE-odds	0.93	0.07	0.07	22.78	0.06	0.12
Commodities	p-value	0.00	0.00	0.22	0.01	0.01	0.95
	Bayesian-p	0.00	0.04	0.78	0.40	0.29	0.34
	BE-odds	1,255.10	4.03	0.05	0.28	0.47	0.37
FX	p-value	0.00	0.20	0.08	0.00	1.00	0.47
	Bayesian-p	0.17	0.78	0.68	0.17	1.00	0.79
	BE-odds	0.93	0.05	0.09	0.93	0.00	0.05
Multi Asset	p-value	0.00	0.00	0.00	0.00	0.03	0.03
	Bayesian-p	0.00	0.00	0.13	0.00	0.51	0.53
	BE-odds	>9,999	105.45	1.26	>9,999	0.18	0.17

Table IV: Historical performance of global return factors: 1800 - 2016

The table summarizes the historical performance of the global return factors. Shown per factor per asset class are the historical annualized Sharpe ratio. The sample starts in January 1800 and ends December 2016 and is at the monthly frequency. Panel A shows the results for the 'new sample' periods (1800-1980 and 2012-2016), Panel B shows the full sample results (1800-2016). Covered are equity indices ("Equities"), 10-year maturity government bond indices ("Bonds"), commodities ("Commodities"), currencies ("FX"), and their equally weighted combination across the four asset classes ("Multi Asset"). Numbers in parentheses indicate t-values. Asterisks are used to indicate significance at a 10% (*), 5% (**) or 1% (***) level.

Panel A: 1800-1980 and 2012-2016

Sharpe ratio	Trend	Momentum	Value	Carry	Seasonality	BAB
Equities	0.75***	0.37***	0.27***	0.68***	0.60***	0.34***
	(11.04)	(5.43)	(3.41)	(9.90)	(8.79)	(4.99)
Bonds	0.49***	0.31***	0.28***	0.44***	0.73***	-0.02
	(7.22)	(4.56)	(3.37)	(6.48)	(10.73)	(-0.29)
Commodities	0.23***	-0.07	0.50***	0.36**	0.85***	-0.06
	(3.38)	(-0.99)	(7.07)	(2.52)	(12.01)	(-0.85)
FX	0.60***	0.32***	0.03	0.52***	0.30***	-0.04
	(8.83)	(4.71)	(0.36)	(7.65)	(4.41)	(-0.59)
Multi Asset	0.98***	0.47***	0.52***	0.91***	1.21***	0.13*
	(14.43)	(6.92)	(7.36)	(13.40)	(17.79)	(1.91)

Panel B: 1800-2016

Sharpe ratio	Trend	Momentum	Value	Carry	Seasonality	BAB
Equities	0.78***	0.43***	0.30***	0.75***	0.61***	0.37***
	(11.49)	(6.31)	(3.79)	(10.91)	(8.94)	(5.43)
Bonds	0.54***	0.33***	0.32***	0.50***	0.69***	0.03
	(7.95)	(4.85)	(3.85)	(7.36)	(10.14)	(0.44)
Commodities	0.35***	0.05	0.49***	0.49***	0.85***	-0.05
	(5.15)	(0.71)	(6.93)	(3.42)	(12.01)	(-0.71)
FX	0.64***	0.33***	0.09	0.56***	0.29***	0.00
	(9.42)	(4.86)	(1.09)	(8.24)	(4.26)	(0.00)
Multi Asset	1.07***	0.55***	0.56***	1.02***	1.19***	0.19***
	(15.76)	(8.09)	(7.92)	(15.02)	(17.49)	(2.79)

Table V: Statistical perspectives on global return factors: 1800 - 2016

The table summarizes various statistical perspectives on the historical performance of the global return factors. Shown per factor per asset class are the historical frequentist p-value ("p-value"), Bayesian p-value using a 4-to-1 prior odds ratio ("Bayesian-p") and break-even prior odds at a 5% confidence level ("BE-odds") of its performance. The sample starts in January 1800 and ends December 2016 and is at the monthly frequency. Panel A shows the results for the 'new sample' periods (1800-1980 and 2012-2016), Panel B shows the full sample results (1800-2016). Covered are equity indices ("Equities"), 10-year maturity government bond indices ("Bonds"), commodities ("Commodities"), currencies ("FX"), and their equally weighted combination across the four asset classes ("Multi Asset").

Panel A: 1800-1980 and 2012-2016

		Trend	Momentum	Value	Carry	Seasonality	BAB
Equities	p-value	0.00	0.00	0.00	0.00	0.00	0.00
	Bayesian-p	0.00	0.00	0.05	0.00	0.00	0.00
	BE-odds	>9,999	>9,999	3.67	>9,999	>9,999	2,021.00
Bonds	p-value	0.00	0.00	0.00	0.00	0.00	0.77
	Bayesian-p	0.00	0.00	0.06	0.00	0.00	0.68
	BE-odds	>9,999	280.39	3.23	>9,999	>9,999	0.09
Commodities	p-value	0.00	0.32	0.00	0.01	0.00	0.40
	Bayesian-p	0.05	0.80	0.00	0.36	0.00	0.80
	BE-odds	3.33	0.05	>9,999	0.33	>9,999	0.05
FX	p-value	0.00	0.00	0.72	0.00	0.00	0.56
	Bayesian-p	0.00	0.00	0.72	0.00	0.00	0.78
	BE-odds	>9,999	546.48	0.07	>9,999	147.25	0.05
Multi Asset	p-value	0.00	0.00	0.00	0.00	0.00	0.06
	Bayesian-p	0.00	0.00	0.00	0.00	0.00	0.64
	BE-odds	>9,999	>9,999	>9,999	>9,999	>9,999	0.11

Panel B: 1800-2016

		Trend	Momentum	Value	Carry	Seasonality	BAB
Equities	p-value	0.00	0.00	0.00	0.00	0.00	0.00
	Bayesian-p	0.00	0.00	0.01	0.00	0.00	0.00
	BE-odds	>9,999	>9,999	13.18	>9,999	>9,999	>9,999
Bonds	p-value	0.00	0.00	0.00	0.00	0.00	0.66
	Bayesian-p	0.00	0.00	0.01	0.00	0.00	0.75
	BE-odds	>9,999	1,040.34	16.36	>9,999	>9,999	0.06
Commodities	p-value	0.00	0.48	0.00	0.00	0.00	0.48
	Bayesian-p	0.00	0.79	0.00	0.05	0.00	0.79
	BE-odds	4,424.81	0.05	>9,999	3.78	>9,999	0.05
FX	p-value	0.00	0.00	0.28	0.00	0.00	1.00
	Bayesian-p	0.00	0.00	0.79	0.00	0.00	1.00
	BE-odds	>9,999	1,090.15	0.05	>9,999	79.16	0.00
Multi Asset	p-value	0.00	0.00	0.00	0.00	0.00	0.01
	Bayesian-p	0.00	0.00	0.00	0.00	0.00	0.23
	BE-odds	>9,999	>9,999	>9,999	>9,999	>9,999	0.63

Table VI: Sub-period performance of global return factors

The table summarizes the historical sub-period performance of the return factors. Shown per factor per asset class is the percentage of rolling 10-year period with positive Sharpe ratios ("$P(SR_{10y}>0)$"). The sample starts in January 1800 and ends December 2016 and is at the monthly frequency. Covered are equity indices ("Equities"), 10-year maturity government bond indices ("Bonds"), commodities ("Commodities"), currencies ("FX"), and their equally weighted combination across the four asset classes ("Multi Asset"). Numbers in parentheses indicate t-values. Asterisks are used to indicate significance at a 10% (*), 5% (**) or 1% (***) level, where in Panel A p-values are based on Newey-West to account for overlapping observations.

$P(SR_{10y}>0)$	Trend	Momentum	Value	Carry	Seasonality	BAB
Equity	98%***	85%***	75%***	95%***	90%***	78%***
Bond	92%***	72%***	83%***	81%***	89%***	34%**
Commodities	73%***	51%	91%***	97%***	99%***	45%
FX	92%***	83%***	61%	80%***	62%	35%
Multi Asset	100%***	91%***	93%***	99%***	100%***	56%

Table VII: Robustness of global return factors: 1800-2016

The table summarizes the robustness of the historical performance of the return factors to strategy construction choices. The factors are tested with the following different implementation methodologies: No liquidity screen, Tercile long-short portfolios (T1-T3), Quartile (Q1-Q4) long-short portfolios ("Quartile portfolios"), no volatility weighting of individual markets or beta scaling of the BAB portfolios ("No volatility or beta scaling"), a 1-month implementation lag ("One month lagged signals"), a quarterly rebalancing period ("Quarterly rebalance period"), and by capping returns of factor returns in a particular market at 50% ("Trimming extreme returns"). We also report the average Sharpe ratio across these various definitions ("Average Sharpe Ratio") and fraction of settings that yield a significant Sharpe ratio ("Fraction"). The sample starts in January 1800 and ends December 2016 and is at the monthly frequency. Shown are the historical annualized Sharpe ratio for each factor per asset class and for the equally weighted multi-asset combination across the four asset classes. Bold numbers indicate significance at the 5% level, while italic numbers are omitted from significance tests.

Factor	Asset Class	Baseline	No liquidity screen	Tercile portfolios	No volatility or beta scaling	One month lagged signals	Quarterly rebalance period	Trimming extreme returns (|50%|)	Average Sharpe Ratio	Fraction
Trend	Eq.	**0.78**	**0.80**	**0.78**	**0.68**	**0.74**	**0.68**	**0.72**	*0.74*	7/7
Trend	Bo.	**0.54**	**0.66**	**0.54**	**0.49**	**0.56**	**0.45**	**0.49**	*0.52*	7/7
Trend	Co.	**0.35**	**0.79**	**0.35**	**0.30**	**0.26**	**0.23**	**0.28**	*0.35*	7/7
Trend	FX	**0.64**	**0.68**	**0.64**	**0.40**	**0.55**	**0.58**	**0.51**	*0.57*	7/7
Momentum	Eq.	**0.43**	**0.52**	**0.41**	**0.40**	**0.36**	**0.35**	**0.39**	*0.40*	7/7
Momentum	Bo.	**0.33**	**0.39**	**0.34**	**0.27**	**0.28**	**0.23**	**0.27**	*0.29*	7/7
Momentum	Co.	0.05	0.12	0.06	0.11	-0.06	-0.08	0.10	*0.04*	0/7
Momentum	FX	**0.33**	**0.54**	**0.32**	**0.25**	**0.39**	**0.30**	**0.29**	*0.32*	7/7
Value	Eq.	**0.30**	**0.34**	**0.31**	**0.31**	**0.32**	**0.29**	**0.32**	*0.30*	7/7
Value	Bo.	**0.32**	**0.35**	**0.30**	**0.37**	**0.25**	**0.22**	**0.35**	*0.29*	7/7
Value	Co.	**0.49**	**0.48**	**0.48**	**0.49**	**0.57**	**0.55**	**0.49**	*0.49*	7/7
Value	FX	0.09	**0.51**	0.09	0.16	0.05	0.16	0.15	*0.15*	1/7
Carry	Eq.	**0.75**	**0.82**	**0.69**	**0.49**	**0.19**	**0.52**	**0.61**	*0.59*	7/7
Carry	Bo.	**0.50**	**0.63**	**0.47**	**0.68**	**0.54**	**0.53**	**0.64**	*0.55*	7/7
Carry	Co.	**0.49**	**0.48**	**0.54**	**0.58**	**0.54**	**0.52**	**0.57**	*0.51*	7/7
Carry	FX	**0.56**	**0.83**	**0.52**	**0.31**	**0.48**	**0.49**	**0.40**	*0.48*	7/7
Seasonality	Eq.	**0.61**	**0.72**	**0.55**	**0.31**	**-0.18**	**0.36**	**0.48**	*0.42*	7/7
Seasonality	Bo.	**0.69**	**1.17**	**0.61**	**0.82**	**-0.14**	**0.48**	**0.83**	*0.65*	7/7
Seasonality	Co.	**0.85**	**0.79**	**0.78**	**0.82**	**0.14**	**0.63**	**0.82**	*0.71*	7/7
Seasonality	FX	**0.29**	**0.46**	**0.31**	0.04	0.04	**0.23**	**0.22**	*0.22*	5/7
BAB	Eq.	**0.37**	**0.38**	**0.35**	-0.11	**0.37**	**0.31**	0.05	*0.22*	5/7
BAB	Bo.	0.03	-0.07	0.05	-0.05	-0.02	0.05	-0.09	*-0.04*	0/7
BAB	Co.	-0.05	**-0.22**	-0.06	0.02	0.01	-0.04	-0.06	*-0.07*	1/7
BAB	FX	0.00	-0.01	-0.01	-0.10	0.02	0.02	-0.03	*-0.03*	0/7
Trend	MA	**1.07**	**1.34**	**1.07**	**0.80**	**0.98**	**0.90**	**0.91**	*1.00*	7/7
Momentum	MA	**0.55**	**0.76**	**0.54**	**0.41**	**0.50**	**0.42**	**0.48**	*0.50*	7/7
Value	MA	**0.56**	**0.76**	**0.55**	**0.63**	**0.55**	**0.58**	**0.59**	*0.57*	7/7
Carry	MA	**1.02**	**1.26**	**0.95**	**0.80**	**0.73**	**0.88**	**0.93**	*0.91*	7/7
Seasonality	MA	**1.19**	**1.50**	**1.10**	**0.98**	**-0.07**	**0.78**	**1.11**	*0.95*	7/7
BAB	MA	**0.19**	0.01	**0.18**	-0.04	**0.21**	**0.18**	-0.06	*0.06*	4/7

Table VIII: Market risk, common variation and global return factors

The table summarizes the market risk and common variation of the global return factors. Panel A shows per factor per asset class the historical annualized Jensen's alpha (in percent) relative to global excess equity, bond, currency and commodity market returns. The alpha's are scaled to a 10% ex-post volatility for ease of comparison. Panel B contains the average pairwise monthly return correlations across the four asset classes ("Factor"), and per asset class or their equally weighted multi-asset aggregation ("Asset class"). Panel C shows the results of spanning tests for each factor return series per asset class on all other factor return series. The return series are scaled to a 10% ex-post volatility for ease of interpretability. The sample starts in January 1800 and ends December 2016 and is at the monthly frequency. Numbers in parentheses indicate t-values. Asterisks are used to indicate significance at a 10% (*), 5% (**) or 1% (***) level.

Panel A: Jensen's alpha

Alpha	Trend	Momentum	Value	Carry	Seasonality	BAB
Equities	7.48***	4.44***	2.36***	7.21***	5.70***	3.32***
	(10.97)	(6.51)	(4.02)	(10.67)	(8.39)	(4.90)
Bonds	5.47***	3.38***	2.16***	4.86***	7.03***	0.36
	(8.00)	(4.96)	(3.87)	(7.14)	(10.32)	(0.54)
Commodities	3.63***	0.51	4.04***	1.09***	7.22***	-0.52
	(5.70)	(0.85)	(6.69)	(3.32)	(11.84)	(-0.89)
FX	6.33***	3.14***	0.68	5.32***	2.73***	-0.27
	(9.27)	(4.60)	(1.22)	(7.83)	(4.02)	(-0.42)
Multi Asset	10.76***	5.58***	4.85***	9.84***	11.69***	1.55**
	(15.74)	(8.17)	(8.00)	(14.45)	(17.13)	(2.31)

Panel B: Average pairwise correlations

	Trend	Momentum	Value	Carry	Seasonality	BAB
Factor	0.05***	0.04**	-0.01	0.00	0.00	0.02
	Equities	**Bonds**	**Commodities**	**FX**	**Multi Asset**	
Asset class	0.08***	0.05***	0.02	0.10***	0.07***	

Panel C: Spanning tests

Intercept (ann.)	Trend	Momentum	Value	Carry	Seasonality	BAB
Equities	3.53***	1.42*	2.25**	5.42***	3.72***	1.07
	(4.93)	(1.91)	(2.46)	(6.68)	(4.63)	(1.39)
Bonds	2.11***	-0.71	4.00***	2.70***	4.66***	1.46*
	(2.99)	(-0.98)	(4.30)	(3.70)	(6.15)	(1.83)
Commodities	2.96***	-0.59	7.07***	3.72**	9.23***	-0.82
	(4.17)	(-0.84)	(8.57)	(2.36)	(10.25)	(-0.90)
FX	4.06***	-1.24**	1.18	3.02***	0.22	-1.18
	(6.88)	(-2.02)	(1.37)	(4.06)	(0.29)	(-1.36)
Multi Asset	6.44***	0.19	7.70***	5.38***	8.32***	0.82
	(10.51)	(0.29)	(9.36)	(7.42)	(11.56)	(1.07)

Table IX: Downside risk and global return factors

The table shows the beta and downside betas of the 24 global return factors and their multi-asset combinations. The downside beta is calculated versus the excess return of the equity market portfolio and uses -1 standard deviation as the threshold. The table is sorted on the difference between downside beta (β^-) and regular CAPM beta (β)and also shows annualized CAPM alpha (α) and DR CAPM alpha (α^-), both in percent. The sample starts in January 1800 and ends December 2016. Bold numbers indicate significant alphas at the 5% level.

Factor	Asset Class	β^-	β	$\beta^- - \beta$	α	α^-
Value	FX	0.01	-0.07	0.08	1.24	0.85
Momentum	Equities	0.04	-0.03	0.07	**4.35**	**4.11**
Carry	FX	0.07	0.01	0.06	**5.53**	**5.31**
Seasonal	Equities	0.10	0.06	0.04	**5.86**	**5.70**
Carry	Commodities	0.03	-0.02	0.04	**4.81**	**4.62**
Momentum	FX	0.06	0.03	0.03	**3.19**	**3.07**
BAB	Commodities	0.14	0.11	0.03	-0.99	-1.11
Seasonal	Bonds	0.00	-0.02	0.02	**7.00**	**6.91**
Value	Equities	-0.04	-0.06	0.02	**3.34**	**3.24**
Carry	Equities	0.06	0.04	0.02	**7.39**	**7.32**
Carry	Bonds	0.07	0.05	0.02	**4.86**	**4.79**
BAB	Equities	0.11	0.09	0.01	**3.39**	**3.33**
Value	Commodities	0.03	0.02	0.01	**5.55**	**5.51**
Seasonal	Commodities	-0.01	-0.02	0.01	**9.30**	**9.26**
Seasonal	FX	0.07	0.06	0.01	**2.68**	**2.64**
Momentum	Commodities	-0.02	-0.02	0.00	0.62	0.61
Trend	FX	0.02	0.03	-0.01	**6.30**	**6.34**
Momentum	Bonds	-0.04	-0.03	-0.01	**3.40**	**3.44**
BAB	Bonds	-0.05	-0.02	-0.03	0.42	0.52
Trend	Commodities	-0.08	-0.05	-0.03	**3.89**	**4.01**
Trend	Bonds	-0.07	-0.04	-0.04	**5.48**	**5.61**
Value	Bonds	-0.07	-0.01	-0.06	**3.20**	**3.48**
BAB	FX	0.05	0.12	-0.07	-0.44	-0.19
Trend	Equities	-0.04	0.06	-0.11	**7.58**	**7.96**
Carry	Multi Asset	0.11	0.05	0.06	**9.95**	**9.73**
Momentum	Multi Asset	0.03	-0.02	0.04	**5.52**	**5.36**
Seasonal	Multi Asset	0.08	0.04	0.04	**11.71**	**11.58**
Value	Multi Asset	-0.04	-0.06	0.02	**6.36**	**6.25**
BAB	Multi Asset	0.12	0.13	-0.01	**1.43**	**1.47**
Trend	Multi Asset	-0.08	0.01	-0.09	**10.68**	**10.99**

Table X: Global return factors in 'good' and 'bad' states

The table summarizes the historical performance of global return factors across various 'good' and 'bad' states based on macroeconomic and market sub-periods. Sub-periods examined are at the annual frequency and include recession versus non-recession, global crisis versus non-crisis, and bear and bull equity markets. Shown are historical (annualized) return per macroeconomic state for each factor-asset class combination and the equally weighted multi-asset combinations across the four asset classes. The column "Dif." contains the differential factor returns between bad and good states. The final row ("Panel") contains the aggregate impact estimated using panel regressions across all global return factors. The panel regression include index fixed-effects and standard errors clustered by date. The sample starts in January 1800 and ends December 2016 and is at the monthly frequency. Bold numbers indicate significance at the 5% level.

		Recession/expansion			Crisis/non-crisis			Bear/bull market		
Factor	Asset Class	Rec.	Exp.	Dif.	Crisis	Non-Crisis	Dif.	Bear	Bull	Dif.
Trend	Eq.	**7.42**	**8.01**	-0.59	**8.81**	**7.49**	1.32	**9.66**	**7.35**	2.31
Trend	Bo.	**5.92**	**5.05**	0.87	**4.51**	**5.61**	-1.11	**4.30**	**5.61**	-1.31
Trend	Co.	**5.41**	**2.68**	2.74	**3.27**	**3.83**	-0.56	**6.40**	**2.98**	3.42
Trend	FX	**4.91**	**7.18**	-2.27	**7.59**	**6.03**	1.55	**8.94**	**5.78**	3.16
Momentum	Eq.	**4.01**	**4.37**	-0.36	**4.68**	**4.11**	0.57	**4.72**	**4.13**	0.60
Momentum	Bo.	**3.13**	**3.37**	-0.25	1.59	**3.83**	-2.23	**3.51**	**3.23**	0.28
Momentum	Co.	0.41	0.52	-0.11	**-2.83**	**1.82**	**-4.65**	1.51	0.20	1.31
Momentum	FX	1.74	**4.12**	-2.38	2.58	**3.53**	-0.94	**3.42**	**3.27**	0.15
Value	Eq.	**5.81**	1.07	**4.74**	**3.60**	**2.83**	0.77	**7.25**	**2.05**	**5.20**
Value	Bo.	2.31	**3.71**	-1.40	1.33	**3.96**	-2.63	**4.27**	**2.84**	1.42
Value	Co.	**5.09**	**5.59**	-0.50	**7.72**	**4.42**	3.30	**5.51**	**5.36**	0.15
Value	FX	0.69	1.00	-0.31	-0.02	**1.27**	-1.30	-0.64	1.27	-1.91
Carry	Eq.	**9.48**	**6.53**	**2.95**	**7.58**	**7.54**	0.04	**8.53**	**7.31**	1.23
Carry	Bo.	**7.08**	**3.98**	**3.10**	**4.52**	**5.21**	-0.68	**6.15**	**4.77**	1.37
Carry	Co.	**7.06**	3.14	3.93	-0.48	**5.91**	-6.40	**8.49**	**3.52**	4.97
Carry	FX	**4.73**	6.01	-1.28	1.88	**6.72**	**-4.83**	**6.42**	**5.36**	1.06
Seasonality	Eq.	**7.31**	5.41	1.90	**5.22**	**6.33**	-1.11	**4.57**	**6.43**	-1.86
Seasonality	Bo.	**8.10**	6.30	1.80	**5.96**	**7.22**	-1.26	**6.58**	**7.00**	-0.42
Seasonality	Co.	**11.02**	7.97	3.05	**7.52**	**9.90**	-2.38	**8.78**	**9.32**	-0.54
Seasonality	FX	1.84	3.47	-1.63	0.92	**3.53**	-2.61	**3.35**	**2.80**	0.55
BAB	Eq.	**2.32**	4.46	-2.15	**4.05**	**3.63**	0.42	1.87	**4.19**	-2.32
BAB	Bo.	0.70	0.16	0.55	1.57	-0.05	1.61	-0.03	0.43	-0.46
BAB	Co.	**-6.46**	3.52	**-9.98**	-2.14	0.14	-2.28	**-7.74**	1.36	**-9.10**
BAB	FX	-1.65	0.84	-2.49	2.18	-0.70	2.88	-1.14	0.23	-1.37
Trend	MA	**10.87**	**10.61**	0.26	**11.21**	**10.54**	0.67	**13.29**	**10.06**	3.23
Momentum	MA	**4.37**	**6.02**	-1.65	**3.12**	**6.19**	-3.08	**6.31**	**5.24**	1.07
Value	MA	**6.80**	**5.61**	1.19	**6.21**	**6.04**	0.17	**7.64**	**5.69**	1.96
Carry	MA	**12.24**	**9.06**	**3.18**	**7.53**	**10.96**	**-3.43**	**12.39**	**9.59**	2.80
Seasonality	MA	**13.91**	**10.79**	**3.12**	**9.73**	**12.52**	-2.79	**11.17**	**12.02**	-0.86
BAB	MA	**-2.79**	**4.32**	**-7.11**	2.32	**1.75**	0.57	**-3.41**	**3.20**	**-6.61**
Panel		**5.00**	**4.95**	0.05	**4.95**	**4.97**	-0.02	**6.33**	**4.62**	**1.71**

Table XI: Macroeconomic risks and global factor returns

The table summarizes the explanatory power of macroeconomic risk for the global return factors using methods outlined in Griffin, Ji and Martin (2003). We explain returns to each global return factor, regressing their returns on global macroeconomic variables. The macroeconomic variables are as in Chen, Roll and Ross (1986): industrial production growth (MP), term premium (UTS), change in expected inflation (DEI), and unexpected inflation (UI). The coefficients and annualized intercept ("Interc. (ann.)") of the regression are shown in the table. We combine the resulting loadings against macroeconomic risks with estimates of risk premiums of these risks (estimated using Fama and MacBeth on all global return factors) to get the predicted return originating from an unconditional macroeconomic risk model ("Predicted (Unc.)"). We also estimate the time-series regressions using a 60-month rolling window and generate one-step ahead forecast of return; we report the average prediction in the column "Predicted (Cond.)". The table further contains the historical average annual return ("Actual") and the differences with predicted returns (i.e. the unexplained return; "Diff."). The sample starts in February 1869 and ends December 2016 and is at the monthly frequency. Bold numbers indicate significance at the 5% level.

Factor	Asset Class	MP	UTS	DEI	UI	Interc. (ann.)	Actual	Predicted (Unc.)	Diff.	Predicted (Cond.)	Diff.
Trend	Eq.	0.00	-0.03	-0.22	-0.03	**9.20**	**9.61**	**4.22**	**5.39**	**-2.20**	**11.72**
Trend	Bo.	**-0.02**	0.02	0.07	0.00	**6.47**	**5.82**	**-2.37**	**8.19**	-0.01	**5.97**
Trend	Co.	0.00	0.09	0.18	-0.03	**3.21**	**4.25**	**-1.93**	**6.18**	**-4.42**	**8.68**
Trend	FX	0.01	-0.05	-0.16	0.00	**6.41**	**6.36**	**2.76**	**3.61**	**4.00**	**2.25**
Momentum	Eq.	0.00	0.00	-0.08	-0.01	**5.66**	**5.73**	**1.07**	**4.66**	**-2.80**	**8.41**
Momentum	Bo.	-0.01	-0.06	-0.07	-0.03	**5.12**	**4.89**	**2.26**	**2.63**	**3.76**	1.22
Momentum	Co.	0.01	0.03	0.09	-0.03	-0.60	0.43	0.29	0.14	**-1.30**	1.85
Momentum	FX	0.00	-0.02	-0.08	0.01	**3.66**	**3.54**	**1.16**	**2.37**	0.90	**2.78**
Value	Eq.	-0.01	0.00	0.07	-0.03	**3.49**	**3.39**	**-0.53**	**3.92**	**1.58**	1.63
Value	Bo.	**-0.04**	0.00	-0.08	0.00	**4.56**	**3.14**	**-1.44**	**4.67**	**3.32**	0.02
Value	Co.	-0.01	**-0.12**	-0.13	**0.03**	**7.16**	**5.37**	**0.89**	**4.48**	**-2.64**	**7.88**
Value	FX	0.02	**-0.13**	0.17	0.01	1.24	0.88	0.21	0.69	**-6.75**	**7.75**
Carry	Eq.	0.01	**-0.19**	0.14	-0.03	**8.91**	**8.46**	**2.62**	**5.84**	-0.68	**9.06**
Carry	Bo.	0.02	0.06	-0.15	-0.02	**4.77**	**6.16**	**2.76**	**3.40**	**-2.34**	**8.64**
Carry	Co.	-0.02	-0.13	0.22	0.00	**6.63**	**4.74**	**-1.51**	**4.42**	**-4.58**	**9.45**
Carry	FX	0.02	**-0.12**	0.11	-0.01	**5.59**	**5.52**	**1.58**	**3.95**	**-2.59**	**8.46**
Seasonality	Eq.	0.01	**-0.25**	**-0.30**	0.00	**6.59**	**5.48**	**7.41**	**-1.94**	**-5.39**	**10.54**
Seasonality	Bo.	0.01	**-0.15**	-0.06	-0.03	**7.37**	**7.16**	**4.26**	**2.90**	-0.63	**7.99**
Seasonality	Co.	**0.03**	-0.06	-0.10	**-0.04**	**8.49**	**9.83**	**5.21**	**4.62**	-1.15	**11.21**
Seasonality	FX	0.01	**-0.12**	-0.11	0.02	**2.84**	**2.16**	**2.63**	-0.51	**-8.33**	**10.69**
BAB	Eq.	0.03	**0.13**	0.07	-0.01	**3.38**	**5.37**	-0.36	**5.73**	1.30	**4.28**
BAB	Bo.	**-0.04**	0.09	0.13	-0.01	1.12	0.40	**-4.47**	**4.87**	**3.46**	**-3.01**
BAB	Co.	0.01	**0.22**	0.06	**0.05**	-1.16	-0.14	**-4.81**	**4.67**	**-3.00**	**2.92**
BAB	FX	0.01	**0.20**	-0.04	**-0.08**	**-2.43**	0.62	**1.58**	-1.11	-1.80	**2.65**
Trend	MA	0.00	0.01	-0.06	-0.02	**11.82**	**12.10**	**1.19**	**10.91**	-0.99	**13.07**
Momentum	MA	0.00	-0.02	-0.07	-0.03	**6.64**	**6.96**	**2.27**	**4.69**	0.36	**6.64**
Value	MA	**-0.02**	**-0.13**	0.02	0.02	**8.44**	**6.51**	-0.52	**7.03**	**-2.45**	**8.93**
Carry	MA	**0.02**	**-0.15**	0.06	-0.03	**10.94**	**11.39**	**3.84**	**7.55**	**-3.84**	**15.38**
Seasonality	MA	**0.03**	**-0.29**	**-0.28**	-0.03	**12.73**	**12.39**	**9.74**	**2.65**	**-7.54**	**20.04**
BAB	MA	0.00	**0.32**	0.12	0.00	0.96	**3.13**	**-5.26**	**8.39**	0.33	**3.07**

Appendix A

A.I: Factor definitions

Time-series momentum ("Trend"). Moskowitz, Ooi, and Pedersen (2012) use a 12-month trend measure for their sample of liquid futures contracts after 1985. We skip the last month on trend investing, as this safeguards against potential liquidity issues that might be especially prominent in the earlier parts of our sample.

Cross-sectional momentum ("Momentum"). We follow Asness, Moskowitz, and Pedersen (2013) and use the 12-month-minus-1-month excess return as momentum measure. We skip the last month on cross-sectional momentum investing, as this safeguards against potential liquidity issues that might be especially prominent in the earlier parts of our sample.

Value ("Value"). For equity, we use the dividend-to-price ratio (D/P), or dividend yield, defined as the past 12-month dividend payment divided by the current price. Other studies typically consider book-to-market value ratios (e.g., Asness, Moskowitz, and Pedersen, 2013), but these are not well available historically. For government bonds, we use the real yield, which is defined as the 10-year bond yield over the past 1-year inflation, as in Asness, Moskowitz, and Pedersen (2013). For currencies, we use an equally-weighted combination of absolute and relative purchasing power parity (PPP), as both are tested in previous studies.[24] Absolute PPP follows Rogoff (1996) and Taylor (2002). Relative PPP is a 5-year reversal of the spot rate corrected for inflation differences defined as in Asness,

[24] Menkhoff, Sarno, Schmeling, Schrimpf (2017) describe more sophisticated currency value measures such as productivity, the quality of export goods, net foreign assets, and output gaps, that we ignore due to limited historical data availability.

Moskowitz, and Pedersen (2013). For commodities, we use the 5-year reversal in spot prices, as defined in Asness, Moskowitz, and Pedersen (2013).

Carry ("Carry"). We use the carry definitions as in Koijen, Moskowitz, Pedersen, and Vrugt (2018). For equity we use the excess implied dividend yield priced into the futures versus spot contract. This effectively captures the implied excess dividend yield of an equity index for the month ahead. We splice these series before the existence of equity futures by means of the following method: We regress the monthly dividend yield implicit in the total versus price return indices on month dummies using the past five years of data to predict the dividend yield for the month ahead, and subsequently subtract the risk-free rate. The average correlation between predicted and actual carry numbers over the period that both are available is 52%. For government bonds, we take the slope of the yield curve defined as the 10-year yield minus the short (3-month) yield and omit the quantitatively smaller roll-down on the interest rate curve because of data limitations. For currencies, we use the short-term yield differential, which we infer from forwards and before their availability from short-term money market rates, and for commodities the slope of the futures curve.

Return seasonality ("Seasonality"). Our definition follows Keloharju, Linnainmaa, and Nyberg (2016): We use the return on an asset in a certain month over the prior 20 years (requiring at least 12-months of data). Assets that did relatively well over a particular month in the past are likely to do relatively well in the same month going forward. For example, in January the monthly equity index seasonal will buy only those equity markets which had the best relative performance in January during the past 20 years and shorts those with lowest relative performance.

Betting-against-beta ("BAB"). This factor postulates that low-beta securities outperform high-beta securities on a beta-adjusted basis (Frazzini and Pedersen, 2014). We test this factor by estimating the betas over a 36-month period (requiring at least 12-

months of data) relative to the global asset class level portfolio return (we refer to Appendix A.II for details about the construction of the global asset class portfolio). The position sizes of each short and long leg are chosen such that the ex-ante betas are the same, so that the excess return contains as little market effect as possible.

A.II Portfolio construction procedure

After obtaining the factor measures per market in each investment universe we construct factor investment portfolios at the end of every month in the following manner. For the trend factor, which is directional in nature, we go long (short) markets in each asset class when the trend measure is positive (negative), following Moskowitz, Ooi, and Pedersen (2012). For the other factors, which are all cross-sectional in nature, we rank the markets in each investment universe based on the factor measure and take a position equal to the rank minus its cross-sectional average (requiring a minimum of two markets to be present). This procedure is similar to that used by Asness, Moskowitz, and Pedersen (2013), Frazzini and Pedersen (2014), and Koijen, Moskowitz, Pedersen, and Vrugt (2018). (By contrast, Keloharju, Linnainmaa, and Nyberg (2016) construct their return seasonality strategy via long the top quintile and short the bottom quintile at each point in time.) Consequently, positions for all factors, except trend, add up to zero at each point in time:

$$\boldsymbol{w_t^i} = \boldsymbol{z_t} \cdot \left(\boldsymbol{Rank}\left(\boldsymbol{S_t^i}\right) - \frac{\boldsymbol{N_t} + \mathbf{1}}{\mathbf{2}}\right),$$

with $\boldsymbol{w_t^i}$ the weight of asset i at time t, $\boldsymbol{S_t^i}$ the factor signal, $\boldsymbol{N_t}$ the number of assets in the cross-section, and $\boldsymbol{z_t}$ a scaling factor to ensure that the portfolio sums to zero.

Next, we size positions in each market in each asset class by its simple 3-year rolling volatility estimate or beta estimate (for BAB only), in the same spirit as Asness, Moskowitz, and Pedersen (2013), Frazzini and Pedersen (2014), and Moskowitz, Ooi, and Pedersen (2012), but fitted to our sample frequency (i.e. monthly data). To prevent undue impact from extremely low volatility estimates (and hence keep the factor strategy robust from an investor perspective), especially in the earlier part of our sample, we floor each volatility (beta) estimate at the maximum of the 10% quantile of volatility (beta) estimates per asset class or 2.5% (0.25), whichever is greater.

We subsequently sum the product of position, sizes, and market returns across markets within an asset class for each date to generate the return on the factor strategy per asset class. We then adjust the position sizes of each of the factor strategies per asset class using a 10-year rolling window such that each factor strategy-asset class combination targets an ex-ante volatility of 10% per annum (an adjustment that implicitly accounts for non-perfect correlations between markets). Again, we floor this estimate to prevent extreme leverage. This approach takes an ex-ante view of portfolio construction, as available in real time. However, our results are not materially different if we simply scale by in-sample ex-post volatility, as in Koijen, Moskowitz, Pedersen and Vrugt (2018).

We rebalance the portfolios each month based on the signals and various volatility estimates. This methodology results in balanced long-short portfolios that are per factor comparable across asset classes, and which will facilitate combining multiple asset classes per factor. Subsequently, we construct such 'Multi Asset' factor portfolios by taking an equally-weighted average of factor portfolios within each asset class and applying a scaling factor equal to the square root of the number of factor series present.[25]

[25] Note that we thereby implicitly assume factor series are uncorrelated across an asset class, an assumption we make for the sake of simplicity. We realize that this assumption is sometimes at odds with the data, but we have verified that this choice does not materially impact our conclusions.

A.III Dataset construction

Financing rates: Our main measure for the financing rates are the 3-month LIBOR rates (sourced from Bloomberg and Datastream), spliced with (in order of usage) eurodollar rates from Datastream, 3-month Treasury bill rates and commercial paper yields from Global Financial Data, short rates from Dimson, Staunton and Marsh, Macrohistory and for the U.S. with data from Jeremy Siegel. When all are unavailable we splice with lagged Treasury-bill returns.

Equities: We source price and return data of equity futures and indices from Bloomberg, Datastream and Global Financial Data. Our primary source is the futures from Bloomberg, with gaps filled in by Datastream data, and spliced before futures inception with index-level data, as in Baltussen, Da, and Van Bekkum (2018). Next, we backfill these data with equity index level data downloaded from Global Financial Data. From the same sources we obtain dividend yields. For carry we use the spot, front futures, and second futures prices. Before we have data on futures, we reconstruct the monthly 'implied carry' as if these markets had listed futures using the regression methodology on the difference between total return and price indices as mentioned in the previous section. The markets we consider are spread around the globe and cover the major developed markets with substantial data history.

Bonds: We source bond futures price and return data from Bloomberg, splice these with bond index-level data from Datastream, backfilled before inception with Global Financial Data. From the same sources we obtain yields, and inflation data, the latter extended where possible with data from Macrohistory. We apply a two-months lag to inflation numbers to mimic their real-time availability. The markets we consider are the major developed bond markets around the globe.

Currencies: Forward and spot prices are primarily from Datastream, spliced with Bloomberg data and Global Financial Data. Before the availability of forward rates, we use short rates from Bloomberg, Datastream, Global Financial Data, Dimson, Staunton and Marsh, and for the U.S., data from Jeremy Siegel. Purchasing power parity data is obtained from the OECD website, and before 1971 with data from Macrohistory. We include the major developed currency markets (or 'G10') in our sample (being USD, GBP, EUR (before 1999: BEF, DEM, ESP, FRF, ITL, and NLG), JPY, CHF, CAD, AUD, NZD, SEK, and NOK), all versus the USD.

Commodities: We source commodity futures price and return data from Bloomberg, spot prices from Bloomberg and Datastream, both spliced with futures data and spot data from Global Financial Data. For carry, we use the front futures and second futures prices. We use the following contracts based on their general usage and liquidity. Futures data available from GFD extents before 1964 for the following six markets: wheat, corn, cotton, cocoa, coffee and copper. Further, we spline the futures data of the front contract with spot data before their inception dates. Due to restrictions on tradability we ignore gold as a speculative asset during the currency gold standard up to the end of the Bretton Woods system (which was effectively a gold standard).

Global asset class portfolios. For equity indices we use the market-value weighted portfolio of equities, spliced before its data existence in 1926 with an equal-weighted portfolio across all equity markets included in the sample. For bonds we use a GDP-weighted global bond portfolio. For currencies, we use an equal-weighted portfolio of all currencies included in our sample versus the US dollar, and for commodities we also use an equal-weighted portfolio of all commodities included in our sample.

Economics: We construct our global recession data from a splicing of the OECD G7 recession indicator from the OECD website (1960-2016), the NBER US recession indicator from the NBER website (1864-1959), and the contraction of real-GDP from Global

Financial Data (1800-1863). Inflation data is described above. We obtain the historical data on crisis periods from Carmen Reinhart and Kenneth Rogoff, using their Banks, Currency, Default, Inflation (BCDI) index, which starts in 1800.[26] International macroeconomic data are from GFD and the OECD. The Chen, Roll and Ross (1986) UTS factor is constructed as the yield on a more than 10-year maturity government bond minus the 3-monthT-bill rate. The factor MP is the log difference in industrial production led by 1-month. Expected inflation and unexpected inflation are calculated following Fama and Gibbons (1984).

[26] http://www.reinhartandrogoff.com/data/

Table A1: A 217 year sample

The table shows the start dates of the individual return series and the value and carry measures of the markets included in the investment universe. Panel A covers equity indices ("Equities"), 10-year maturity government bond indices ("Bonds"), commodities ("Commodities"), and currencies ("FX").

Starting date	Country	Index	Returns	Value	Carry
Equities	U.S.	S&P500	1800-02-28	1801-01-31	1805-01-31
	U.K.	FTSE100	1799-12-31	1923-12-31	1799-12-31
	Germany	DAX	1870-01-31	1869-12-31	1874-12-31
	France	CAC	1856-01-31	1855-12-31	1860-12-31
	Japan	TOPIX	1886-01-31	1886-12-31	1890-12-31
	Australia	ASX200	1882-10-31	1882-10-31	1887-09-30
	Canada	TSE60	1934-01-31	1934-01-31	1938-12-31
	Switzerland	SMI	1966-02-28	1966-01-31	1971-01-31
	Sweden	OMX	1902-01-31	1871-12-31	1906-12-31
	Spain	IBEX	1876-01-31	1875-12-31	1880-12-31
	Italy	MIB	1925-01-31	1925-01-31	1929-12-31
	the Netherlands	AEX	1951-01-31	1969-07-31	1955-12-31
	Hong Kong	HANGSENG	1969-12-31	1972-12-31	1974-12-31
	South Korea	KOSPI200	1962-02-28	1963-01-31	1967-01-31
Bonds	U.S.	10Y bond	1799-12-31	1872-02-29	1799-12-31
	U.K.	10Y bond	1799-12-31	1872-02-29	1799-12-31
	Germany	10Y bond	1815-03-31	1869-03-31	1815-02-28
	France	10Y bond	1800-02-28	1872-02-29	1800-01-31
	Japan	10Y bond	1880-01-31	1872-02-29	1879-12-31
	Australia	10Y bond	1857-06-30	1872-02-29	1857-06-30
	Canada	10Y bond	1934-03-31	1872-02-29	1934-02-28
	Switzerland	10Y bond	1900-01-31	1893-12-31	1893-12-31
	Sweden	10Y bond	1853-10-31	1872-02-29	1854-01-31
	Spain	10Y bond	1881-01-31	1872-02-29	1880-12-31
	Italy	10Y bond	1862-01-31	1872-02-29	1861-12-31
	the Netherlands	10Y bond	1814-03-31	1872-02-29	1814-02-28
	New Zealand	10Y bond	1923-02-28	1915-10-31	1923-01-31
	Norway	10Y bond	1822-04-30	1872-02-29	1822-03-31
	Belgium	10Y bond	1848-09-30	1872-02-29	1848-08-31
	Denmark	10Y bond	1864-03-31	1872-02-29	1864-02-29

Starting date	Country	Index	Returns	Value	Carry
Commodities	-	Soy oil	1911-02-28	1915-07-31	1968-04-30
	-	Corn	1858-02-28	1862-07-31	1968-06-30
	-	Cocoa	1799-12-31	1799-12-31	1967-12-31
	-	Oil (WTI)	1859-10-31	1864-03-31	1983-05-31
	-	Oil (Brent)	1988-10-31	1974-08-31	1989-02-28
	-	Cotton	1799-12-31	1799-12-31	1968-02-29
	-	Gold	1974-01-31	1974-01-31	1975-01-31
	-	Copper	1800-02-28	1804-07-31	1994-10-31
	-	Heating Oil	1967-02-28	1971-07-31	1986-06-30
	-	Gasoline	1986-07-31	1987-07-31	1986-12-31
	-	Coffee	1825-02-28	1829-07-31	1973-02-28
	-	Kansas Wheat	1970-03-31	1983-06-30	1970-02-28
	-	Aluminum	1901-01-31	1905-06-30	1997-08-31
	-	Cattle	1858-02-28	1862-07-31	1965-04-30
	-	Live Hog	1858-02-28	1862-07-31	1987-01-31
	-	Nickel	1901-01-31	1905-06-30	1997-08-31
	-	Zinc	1840-02-29	1844-07-31	1997-08-31
	-	Natural Gas	1890-01-31	1894-06-30	1990-05-31
	-	Soybeans	1913-12-31	1918-05-31	1968-04-30
	-	Sugar	1960-02-29	1964-07-31	1968-02-29
	-	Silver	1799-12-31	1799-12-31	1975-01-31
	-	Soymeal	1929-11-30	1934-04-30	1968-04-30
	-	Wheat	1799-12-31	1799-12-31	1968-06-30
FX	-	AUD/USD	1834-12-31	1871-02-28	1834-11-30
	-	CAD/USD	1934-03-31	1871-02-28	1934-02-28
	-	CHF/USD	1831-01-31	1871-02-28	1830-12-31
	-	DEM/USD	1815-03-31	1871-02-28	1815-02-28
	-	EUR/USD	1999-01-31	1871-02-28	1815-02-28
	-	GBP/USD	1799-12-31	1871-02-28	1799-12-31
	-	JPY/USD	1880-01-31	1871-02-28	1879-12-31
	-	NOK/USD	1819-05-31	1871-02-28	1819-04-30
	-	NZD/USD	1923-02-28	1920-09-30	1923-01-31
	-	SEK/USD	1836-03-31	1871-02-28	1836-02-29
	-	FRF/USD	1800-02-28	1871-02-28	1800-01-31
	-	ITL/USD	1862-01-31	1871-02-28	1861-12-31
	-	ESP/USD	1881-01-31	1871-02-28	1880-12-31
	-	NLG/USD	1814-03-31	1871-02-28	1814-02-28
	-	BEF/USD	1848-09-30	1871-02-28	1848-08-31

Appendix B: Additional results

Table B.1: Robustness of global return factors: replication sample

The table summarizes the robustness of the historical performance of the global return factors over the replication sample to common strategy construction choices made by the original studies. Factors are tested with the following different implementation methodologies: Tercile long-short portfolios (T1-T3), and no volatility weighting of individual markets or beta scaling of the BAB portfolios ("No volatility or beta scaling"). Shown are the historical annualized Sharpe ratios for each factor per asset class and for the equally weighted multi-asset combination across the four asset classes. The sample starts in January 1981 and ends December 2011 and is at the monthly frequency. Bold numbers indicate significance at the 5% level.

Factor	Asset Class	Baseline	Tercile portfolios	No volatility or beta scaling
Trend	Equities	**0.62**	**0.62**	**0.53**
Trend	Bonds	**0.53**	**0.53**	**0.49**
Trend	Commodities	**0.88**	**0.88**	**0.75**
Trend	FX	**0.53**	**0.53**	**0.45**
Momentum	Equities	**0.61**	**0.61**	**0.55**
Momentum	Bonds	0.29	0.35	0.13
Momentum	Commodities	**0.62**	**0.64**	**0.61**
Momentum	FX	0.23	0.20	0.27
Value	Equities	0.34	0.35	**0.47**
Value	Bonds	0.29	0.28	**0.61**
Value	Commodities	0.22	0.22	0.16
Value	FX	0.32	**0.39**	**0.43**
Carry	Equities	**0.87**	**0.72**	**0.82**
Carry	Bonds	**0.71**	**0.71**	**0.66**
Carry	Commodities	**0.44**	**0.53**	**0.58**
Carry	FX	**0.53**	**0.50**	**0.61**
Seasonality	Equities	**0.36**	0.31	**0.37**
Seasonality	Bonds	-0.09	-0.14	-0.14
Seasonality	Commodities	**0.48**	**0.44**	**0.42**
Seasonality	FX	0.00	0.04	-0.04
BAB	Equities	**0.37**	**0.36**	-0.10
BAB	Bonds	**0.36**	**0.36**	-0.11
BAB	Commodities	0.01	0.03	-0.04
BAB	FX	0.13	0.13	0.06
Trend	Multi Asset	**1.09**	**1.09**	**0.93**
Momentum	Multi Asset	**0.78**	**0.80**	**0.78**
Value	Multi Asset	**0.55**	**0.58**	**0.61**
Carry	Multi Asset	**1.15**	**1.12**	**1.18**
Seasonality	Multi Asset	**0.40**	0.35	**0.49**
BAB	Multi Asset	**0.39**	**0.39**	-0.10

Table B.2: Common variation per global return factor

The table shows the individual correlation coefficients of the global return factors. Shown are the monthly return correlations between each factor-asset class series. The sample starts in January 1800 and ends December 2016 and is at the monthly frequency.

		Trend				Momentum				Value				Carry				Seasonal				BAB			
		EQ	BND	COM	FX	EQ	BND	COM	FX	EQ	BND	COM	FX	EQ	BND	COM	FX	EQ	BND	COM	FX	EQ	BND	COM	FX
Trend	Equity		0.13	0.07	0.04	0.37	0.09	0.03	0.04	-0.08	0.05	-0.03	0.00	0.08	0.00	0.07	0.01	0.11	0.01	0.02	0.01	0.24	0.04	0.00	0.05
	Bond	0.13		0.01	0.03	0.11	0.46	0.00	0.01	-0.03	-0.16	0.01	-0.01	0.02	0.24	0.01	0.03	-0.03	0.23	0.02	0.01	0.01	-0.05	0.01	-0.02
	Commo	0.07	0.01		0.04	0.04	0.05	0.60	0.02	-0.03	0.01	-0.29	-0.02	0.02	-0.01	0.21	-0.04	-0.02	-0.01	0.04	-0.05	-0.01	0.02	0.03	-0.02
	FX	0.04	0.03	0.04		0.02	0.05	-0.01	0.65	-0.01	0.00	-0.03	-0.30	0.04	0.02	-0.03	0.12	0.01	0.00	-0.04	0.33	0.00	-0.03	0.03	0.10
Momentum	Equity	0.37	0.11	0.04	0.02		0.12	0.03	-0.01	-0.09	0.06	-0.04	-0.01	0.05	-0.01	0.01	0.00	-0.05	0.02	0.03	-0.01	-0.12	0.02	0.02	0.00
	Bond	0.09	0.46	0.05	0.05	0.12		0.06	0.04	-0.07	0.07	-0.03	-0.06	-0.01	0.18	0.03	0.01	0.00	0.16	0.04	0.01	0.02	-0.03	0.05	-0.02
	Commo	0.03	0.00	0.60	-0.01	0.03	0.06		-0.01	-0.02	0.03	-0.42	-0.01	0.02	-0.03	0.33	0.01	-0.01	0.00	0.01	-0.04	-0.02	0.02	0.05	-0.01
	FX	0.04	0.01	0.02	0.65	-0.01	0.04	-0.01		-0.03	0.01	-0.01	-0.31	0.02	0.01	0.03	0.12	0.00	-0.01	0.01	0.19	0.00	0.00	0.04	0.03
Value	Equity	-0.08	-0.03	-0.03	-0.01	-0.09	-0.07	-0.02	-0.03		-0.05	-0.01	-0.01	0.18	0.06	0.01	-0.01	0.11	0.01	-0.02	0.00	0.12	-0.01	-0.03	0.03
	Bond	0.05	-0.16	0.01	0.00	0.06	0.07	0.03	0.01	-0.05		-0.02	0.01	-0.04	-0.30	0.08	0.01	-0.06	-0.14	0.02	0.00	-0.02	0.18	0.03	-0.01
	Commo	-0.03	0.01	-0.29	-0.03	-0.04	-0.03	-0.42	-0.01	-0.01	-0.02		0.03	-0.04	0.02	-0.24	0.02	-0.02	0.01	-0.04	0.02	-0.02	0.00	0.00	0.04
	FX	0.00	-0.01	-0.02	-0.30	-0.01	-0.06	-0.01	-0.31	-0.01	0.01	0.03		0.00	0.00	-0.06	0.36	0.00	-0.02	-0.01	-0.02	0.01	0.01	-0.02	-0.05
Carry	Equity	0.08	0.02	0.02	0.04	0.05	-0.01	0.02	0.02	0.18	-0.04	-0.04	0.00		0.03	-0.06	0.02	0.12	0.05	0.02	0.01	0.04	0.01	-0.05	-0.02
	Bond	0.00	0.24	-0.01	0.02	-0.01	0.18	-0.03	0.01	0.06	-0.30	0.02	0.00	0.03		0.00	0.04	0.05	0.35	-0.01	-0.01	0.04	-0.22	0.04	-0.04
	Commo	0.07	0.01	0.21	-0.03	0.01	0.03	0.33	0.03	0.01	0.08	-0.24	-0.06	-0.06	0.00		-0.01	0.08	-0.03	-0.02	-0.02	0.08	0.02	0.04	0.07
	FX	0.01	0.03	-0.04	0.12	0.00	0.01	0.01	0.12	-0.01	0.01	0.02	0.36	0.02	0.04	-0.01		0.01	0.03	-0.01	0.26	0.00	-0.03	0.02	-0.05
Seasonal	Equity	0.11	-0.03	-0.02	0.01	-0.05	0.00	-0.01	0.00	0.11	-0.06	-0.02	0.00	0.12	0.05	0.08	0.01		0.02	0.03	0.01	0.17	-0.02	-0.03	0.01
	Bond	0.01	0.23	-0.01	0.00	0.02	0.16	0.00	-0.01	0.01	-0.14	0.01	-0.02	0.05	0.35	-0.03	0.03	0.02		0.01	-0.04	0.02	-0.20	0.02	0.00
	Commo	0.02	0.02	0.04	-0.04	0.03	0.04	0.01	0.01	-0.02	0.02	-0.04	-0.01	0.02	-0.01	-0.02	-0.01	0.03	0.01		-0.01	0.00	-0.01	-0.01	-0.01
	FX	0.01	0.01	-0.05	0.33	-0.01	0.01	-0.04	0.19	0.00	0.00	0.02	-0.02	0.01	-0.01	-0.02	0.26	0.01	-0.04	-0.01		-0.02	-0.03	0.05	0.09
BAB	Equity	0.24	0.01	-0.01	0.00	-0.12	0.02	-0.02	0.00	0.12	-0.02	-0.02	0.01	0.04	0.04	0.08	0.00	0.17	0.02	0.00	-0.02		0.01	0.02	0.03
	Bond	0.04	-0.05	0.02	-0.03	0.02	-0.03	0.02	0.00	-0.01	0.18	0.00	0.01	0.01	-0.22	0.02	-0.03	-0.02	-0.20	-0.01	-0.03	0.01		-0.01	0.01
	Commo	0.00	0.01	0.03	0.03	0.02	0.05	0.05	0.04	-0.03	0.03	0.00	-0.02	-0.05	0.04	0.04	0.02	-0.03	0.02	-0.01	0.05	0.02	-0.01		0.04
	FX	0.05	-0.02	-0.02	0.10	0.00	-0.02	-0.01	0.03	0.03	-0.01	0.04	-0.05	-0.02	-0.04	0.07	-0.05	0.01	0.00	-0.01	0.09	0.03	0.01	0.04	

Table B.3: Different downside beta definitions of global return factors

The table shows the downside betas of the multi-asset combinations of the global return factors using deeper tail thresholds and equity or bond markets. Panel A shows downside beta versus excess return of the equity market portfolio. Panel B shows downside beta versus excess return of the bond market portfolio. We depict downside betas for thresholds of 0, -1, -2 and -3 standard deviations away from zero and the number of observations below the downside risk threshold. The two panels are sorted on the -3 standard deviation threshold ($\beta^{-3stdev}$). The sample starts in January 1800 and ends December 2016 and is at the monthly frequency.

Panel A: Equity market downside risk

Factor	Asset Class	β^{-0}	$\beta^{-1stdev}$	$\beta^{-2stdev}$	$\beta^{-3stdev}$
Carry	Multi Asset	0.09	0.11	0.18	0.24
BAB	Multi Asset	0.11	0.12	0.13	0.12
Seasonal	Multi Asset	0.06	0.08	0.10	0.11
Value	Multi Asset	-0.04	-0.04	0.03	0.01
Momentum	Multi Asset	0.01	0.03	0.01	0.01
Trend	Multi Asset	-0.05	-0.08	-0.11	-0.08
# observation < equity threshold		1104	218	74	30

Panel B: Bond market downside risk

Factor	Asset Class	β^{-0}	$\beta^{-1stdev}$	$\beta^{-2stdev}$	$\beta^{-3stdev}$
Value	Multi Asset	-0.01	0.02	0.04	0.04
Seasonal	Multi Asset	0.03	0.05	0.04	0.03
Momentum	Multi Asset	0.05	0.07	0.03	0.03
Trend	Multi Asset	0.03	0.02	0.00	0.02
Carry	Multi Asset	0.03	0.04	-0.01	0.00
BAB	Multi Asset	-0.03	-0.03	-0.08	-0.11
# observations < bond threshold		1018	161	44	20