Beyond carry and momentum in government bonds

Jérôme Gava* William Lefebvre † Julien Turc ‡

September 11, 2019

Abstract

This article revisits recent literature on factor investing in government bonds, in particular regarding the definition of value and defensive investing. Using techniques derived from machine learning, the authors identify the key drivers of government bond futures and the groups of factors that are most genuinely relevant. Beyond carry and momentum, they propose an approach to defensive investing that considers the safe-haven nature of government bonds. These two main styles may be complemented by value and a reversal factor in order to achieve returns independently from broad movements in interest rates.

Keywords: Factor and style investing, fixed income, macroeconomy, macro-finance, seasonality, mean-reversion, machine learning, variables selection, panel regression, clustering, covariance selection, selection bias under multiple testing

*Jérôme Gava, PhD, is a senior researcher in the QIS Lab at BNP Paribas, Paris and visiting researcher at the department of economics of Ecole Polytechnique, Palaiseau.

†William Lefebvre is a researcher in the QIS Lab at BNP Paribas, Paris and PhD student at the LPSM, Paris Diderot University - Paris 7.

‡Corresponding author: Julien Turc (julien.turc@bnpparibas.com) is head of the QIS Lab at BNP Paribas, Paris and visiting researcher at the department of economics of Ecole Polytechnique, Palaiseau.

Contents

Editorial **3**

Taking a fresh look at styles investing **5**

50 years of data 5

A broad overview of styles in government bonds 7

The drivers of government bond futures **10**

Seasonality and mean reversion 10

Selecting the right variable 12

Looking for evidence on factors returns **17**

Style investing in practice 17

Clusters and extended styles 22

Which clusters are genuinely relevant? 25

Conclusion **27**

Appendices **32**

FACTOR investing in fixed income has been the subject of much recent research. Factors, or styles, hold the promise of positive long-term returns with little directional risk (Asness et al. 2015). Applied to fixed income, such factors would allow portfolio managers to outperform their benchmarks in the long run and asset allocators to opportunistically reduce average interest rates risk whilst maintaining expected returns.

However, bonds are not equities and this point is particularly clear in the case of government bonds, which are the focus of this study. In an echo of a seminal paper by Asness, Moskowitz, and Pedersen (2013) ("Value and Momentum Everywhere"), Bhansali et al. (2015) share the perspective of the fixed income investor and stress the ubiquity of carry and trends across asset classes. While these authors actually discuss directional exposures, the stage is set for an exchange of views between the proponents of factor investing (which originated in stock picking) and fixed income specialists.

Fattouche (2018) presents evidence on carry, value and momentum in government bonds. Following Asness et al. (2015), Brooks, Palhares, and Richardson (2018) also consider a defensive factor. Whilst there is little controversy regarding the nature of carry and momentum, defining value and defensive investing in fixed income is far from straightforward.

Value involves setting market prices against a fundamental anchor. Surprising as it may be for an equity analyst, there is no agreement in the fixed income community as to the right rate of return to apply for a fundamental assessment of value. Relative value analysis is a more common approach and involves averages across sectors or buckets. Research in the factors community variously considered past interest rates, inflation and growth as possible anchors. Another stream of research (Rudebusch 2010), largely driven by central banks, relies on macroeconomic models in order to link certain fundamental variables with the shape of the curve.

Defensive investing in fixed income has been discussed from the perspective of adjusting duration (Leote de Carvalho et al. 2014, Brooks, Palhares, and Richardson 2018). This is a fair point, but managing duration can been seen as taking a view on the steepness of the curve, which is also a key market factor among bonds of various maturities (Litterman and Scheinkman 1991). Some authors (Brooks and Moskowitz 2017) even look to invest into the three main components of the curve by following styles. Defensive investing is not one of them.

Even in the stock market, defining styles and factors is a vexed issue that has been hotly debated. Econometric models (see Fama and French 2015) played a key role in this debate, even though recent research on multiple asset classes has focused on a selected set of commonly accepted factors. In our investigation of government bond futures, we broaden the range of possible factors and look to inductively select the most relevant ones.

Of course, this approach exposes us to false discoveries and overfitting, subjects that have been much discussed recently. We follow Arnott, Harvey, and Markowitz (2019) in establishing an *ex ante* economic foundation, keeping track of simulations and considering trading costs and constraints. As interest rates have fallen almost continuously over much of the readily available data, we strive to eliminate any market biases and extend the data set to the 70's and 80's. We then use techniques pioneered by Fabozzi and López de Prado

(2018) in order to adjust findings for the selection bias under multiple testing. We also introduce an indicator to measure the robustness of factors across countries.

Going back to the roots of factor investing, we complement the factors mining approach by an analysis of statistical drivers. Rather than regressing the returns of bonds versus those of factor portfolios, we follow Brooks and Moskowitz (2017) in considering leading indicators directly. Doing so allows us to relate the bonds of a given country with local indicators and opens the way for an investigation of time series and cross section patterns. We use a machine learning algorithm for selecting the best combination of variables and contribute to a growing body of research concerning the application of machine learning in fixed income (Ludvigson and Ng 2009, Bianchi, Büchner, and Tamoni 2019).

When all is said and done, we find evidence for 4 main investment styles. Carry is ubiquitous and strongly linked with momentum. Defensive investing also passes statistical tests after adjustment for multiple testing. That bonds can provide a useful safe haven shall come as no surprise and has been recently stressed by Baz, Sapra, and Ramirez (2019). We contribute to the literature on factor investing by showing that monitoring a combination of economic and technical variables is a sensible approach to defensive investing and a good complement to carry.

Our simple approach to value and reversal may provide further diversification, although these two factors do not look as statistically relevant as carry and momentum. These approaches can be further refined, by considering forward looking information (Brooks, Palhares, and Richardson 2018) and adjusting for changes in central banks policy that took place in the early '90s and over the last decade.

The first part of this study presents the investment universe and the data set. The second part provides information on time patterns and the key drivers of returns in the time series and cross section. The third part focuses on cross-sectional factor strategies and involves techniques from machine learning for identifying the main styles.

Taking a fresh look at styles investing

50 years of data

A number of studies, especially those that investigate the link between macroeconomics and financial data (Rudebusch 2010), focus on the US, where most historical data is available. A more recent strand of literature looks for evidence on alternative factors across various countries. As historical depth varies across countries, some authors start their investigations in the 1990's, when data is available for all countries. Others increase the data set over time. Financial instruments also vary.

Table 1: Data sets in selected studies

Authors	**Countries**	**Start point**	**I**[‡]	**Weights**[*]	**TC**[§]
Cochrane and Piazzesi (2005)	US	1964	B	-	-
Asness, Moskowitz, and Pedersen (2013)	10 developed countries	1982[†]	B	E or R	No
Baz et al. (2015)	G10 + 4 EM	1990	S	E	No
Bhansali et al. (2015)	AU, DE, JP, GB, US	1972[†]	B, F	-	-
Brooks and Moskowitz (2017)	AU, CA, DE, JP, SE, GB, US	1971[†]	B	D + R	No
Baltas (2017)	AU, CA, CH, DE, JP, NZ, SE, US	1982[†]	F	R	No
Fattouche (2018)	AU, CA, DE, JP, GB, US	1960	B, F	E	Yes
Brooks, Palhares, and Richardson (2018)	13 developed countries	1996	B	E	No

[†] *Earliest starting point. Time series are added as data becomes available, with the latest starting point falling between 1990 and 2002.*
[‡] *Instruments. B stands for bonds, F for futures and S for swaps.*
[*] *Portfolio weights. E stands for equal, D for duration-adjusted, R for ranks-adjusted.*
[§] *Transaction costs*

Our first data set starts in 1992 with daily quotes. We consider 10 year government bond futures for 6 countries, namely Australia, Canada, Germany, Japan, the UK and the US. These markets were selected for their high level of liquidity. The list includes only one country from the Euro area. Credit risk linked to a country not controlling monetary policy falls beyond the scope of this study. When simulating strategies, we apply transaction costs that are in line with current market levels and add a cost for running the strategies.

Given the lack of information about rising rates and inflation in these data, it is worth considering events from the Great Inflation, however imperfect the data may be (Bhansali et al. 2015). For the purposes of the statistical analysis, it is also useful to consider data that start on the same date for all countries. Our second data set starts in 1969 for the same list of countries and does not include Japan. The model we use for generating the data is detailed in appendix A and we consider monthly data until 1992.

The first 'Great Moderation' data set is the most relevant when estimating the return of investment strategies, at least under the current inflation regime. We view the second 'Inflation and Moderation' set as an additional source of information on what could happen if inflation and rates were to rise again.

Although cross-sectional factors are our main focus, we also investigate statistical patterns in the time series, following Brooks and Moskowitz (2017). Comparing both sheds more light on the peculiar dynamics at stake in a market-neutral portfolio. While some authors measure cross-sectional returns by ranking them (Asness, Moskowitz, and Pedersen 2013) or removing time fixed effects (Brooks and Moskowitz 2017), we adjust each bond future for its beta against a market factor[1].

The alpha of a bond future represents its outperformance against the market, after adjusting for correlations. It is an interesting source of information on individual bonds and factor portfolios [2]. The cross-sectional standard deviation of the alphas is 0.25% on average in both data sets.

Table 2: Winners and losers among government bond futures

	Great Moderation						**Inflation and Moderation**				
	AU	**CA**	**DE**	**GB**	**US**	**JP**	**AU**	**CA**	**DE**	**GB**	**US**
Bond futures											
Annual return (%)	3.8	3.4	4.1	3.3	3.4	3.2	0.7	2.0	2.8	2.2	2.3
MDD (%)	-20	-16	-11	-17	-14	-9	-64	-40	-34	-31	-52
Annualised vol (%)	7.5	5.9	5.3	6.4	5.8	3.9	7.6	5.9	6.1	7.6	7.0
Information Ratio†	0.51	0.59	0.77	0.51	0.58	0.82	0.09	0.34	0.47	0.29	0.33
Market exposure‡(%)	51	76	79	77	76	29	52	69	67	67	62
Alphas											
Annual return (%)	-0.2	-1.0	0.6	-1.2	-0.8	2.1	-0.9	-0.7	1.5	0.0	-0.5
Max. drawdown (%)	-33	-27	-8	-36	-23	-8	-52	-40	-15	-44	-40
Annualised vol (%)	6.5	3.8	3.2	4.1	3.8	3.7	6.5	4.2	4.5	5.6	5.6
Information Ratio†	-0.03	-0.26	0.19	-0.3	-0.21	0.55	-0.14	-0.17	0.33	0	-0.08

† *Return/ vol*

‡ *Market exposure is the correlation between bond futures and the market factor, considering weekly returns in the first data set and monthly returns for the second one.*

Sources: BNP Paribas, Bloomberg, Federal Reserve of St Louis, Bundesbank

Government bond futures have performed greatly since the early 1990's, a bit less so when considering the earlier period. Relative returns vary greatly across countries, with futures in Germany by far outperforming those in Australia, Canada and the US. Since 1992, Japanese futures have performed best and British futures worst. Japanese and, to a lesser extent, Australian bonds are also less correlated with the market factor than bonds of other countries.

1. For each data set, we build a market factor that allocates across all countries with an equal weight, adjusted daily. The beta of each bond future is estimated by regressing future returns against the market factor, considering overlapping weekly changes and a window of 1y for the first data set and monthly changes with a window of 3y for the second data set.

2. We consider the resulting alpha as a virtual trading instrument and an interesting source of information. and apply the same transaction costs as for bond futures. Doing so understates the cost of hedging for a given country, but avoids duplicating costs when analyzing a market-neutral portfolio.

A broad overview of styles in government bonds

Following Asness et al. (2015), we look for evidence of four main investment styles, namely carry, momentum, value and defensive. Brooks, Palhares, and Richardson (2018) follow a similar approach in fixed income. We consider three additional styles, each of which is related to one of the four main styles.

The remainder of this study presents statistical evidence on these factors and which indicator is most relevant in order to capture them. The table in appendix C sums up the list of indicators for each style. A number of indicators involve parameters, most of the time in order to define a moving average or a change. We refer to such particular versions of an indicator as variables.

In the second part of this study, we select a set of variables and identify those that lead the returns of government bonds. In the third part, we associate each possible variable to a factor portfolio and identify clusters of factors. In the event, we end up with a list of extended styles that differs from the four styles identified by Asness et al. 2015 across asset classes.

It is now necessary to define the indicators.

Carry and curve: The spread between the yield of a government bond and the money market is a key source of carry (Brooks, Palhares, and Richardson 2018), a term we refer to as 'simple carry'. Some authors (Bhansali et al. 2015, Koijen et al. 2018) also consider the additional benefits of rolling down the curve. Given the lack of relevant information, we rely on simple carry for the Inflation and Moderation data set and include roll-down[3] for the Great Moderation.

Carry is a form of reward for holding a long-dated security. In the model of Heath, Jarrow, and Morton (1992), long-term rates y can be split into three components:

$$y(0,T) = \mathbb{E}\left[\int_0^T r(t)\frac{dt}{T}\right] + \mathbb{E}\left[\int_0^T \lambda(t)\sigma(t,T)\frac{dt}{T}\right] - \frac{1}{2}\mathbb{E}\left[\int_0^T \sigma(t,T)^2\frac{dt}{T}\right] \quad (1)$$

where r are short-term rates in the future, λ the market price of risk and σ a volatility[4]. Steepness in the curve reflects expectations about short-term rates, aversion to risk and volatility. If investors were not averse to risk and expected rates to remain stable, the curve would be slightly inverted and negatively convex due to a convexity premium. But rate expectations may not be constant and aversion to risk implies a steeper curve.

The investor who is not afraid of risk can hope to monetize the price of risk, all the more so if volatility is high. She may also disagree with expectations that are embedded in the curve, or conversely view them as a useful source of information.

3. Simple carry is the yield of a 10y government bond versus the local 3m interbank rate. Our measure of roll-down is based on a linear interpolation between 5y and 10y bonds.

4. Technically, $y(t,T)$ and $\sigma(t,T)$ are associated at time t to a zero-coupon that expires at time T. A common specification is $\sigma(t,T) = \sigma \cdot (1 - \exp(-a \cdot (T-t)))/a$, where a represents the mean-reversion of short-term rates.

Steepness and curvature are principal components of bond returns, as shown by Litterman and Scheinkman (1991). Brooks and Moskowitz (2017) compare the impact of these curve factors with that of other alternative factors. Fama and Bliss (1987), Cochrane and Piazzesi (2005) forecast the returns of US Treasury bonds using forward rates. Kessler and Scherer (2009) extend this approach to other countries. Simple carry is an indicator of steepness and roll-down captures the shape on the long end. Without delving further into this topic, we include information on the short end of the curve and a convexity factor among our indicators[5]. We do so for the Great Moderation data set, no data being previously available for all countries.

The short-end of the curve may bring valuable insight on the expected path for monetary policy. A highly negative convexity reflects high expected volatility and more prosaically a low roll-down combined with monetary tightening.

Momentum: It is classically defined as the returns from the past 12 months (Brooks, Palhares, and Richardson 2018). Bhansali et al. (2015) consider binary signals, Baz et al. (2015) average cross-over signals from various horizons and Baltas and Kosowski (2017) review different ways to measure trends efficiently. In the second part of this study, we follow Asness, Moskowitz, and Pedersen (2013) in skipping the most recent month and smooth the data by averaging the indicator over the past 21 business days. These two parameters are allowed to vary in the third part.

It shall be noted that a major policy shift occurred in the early 1990's, when central banks adopted inflation targeting (Bernanke and Mishkin (1997)) and focused on short-term interest rates as operating objectives (Borio 1997). With quantitative easing, central banks have played an increasing role in long-term interest rates. These changes are likely to have had a major impact on momentum patterns. The wild fluctuations that occurred during the Great Inflation and the subsequent decrease in rates were not seen again and may not be as long as inflation expectations remain anchored.

Value, reversal and fundamentals: There is less consensus in the literature as to the value factor. Asness, Moskowitz, and Pedersen (2013), followed by Dorsten, Davis, and Rennison (2016), define value as the change in yields over the last 5 years, a classic reversal signal.

Arguably, a measure of value is supposed to set a market price against a fundamental anchor. Asness et al. (2015) and Brooks and Moskowitz (2017) compare nominal yields with inflation expectations or past CPI inflation, based on availability. Fattouche (2018) compares nominal yields with past inflation and economic growth [6].

Baz et al. (2015) justify this choice through an intertemporal macroeconomic model. Their equilibrium real rate ensures that consumers are indifferent between spending now or in 10 years. Equating consumption to output and neglecting uncertainty, real rates would fluctuate around economic growth. This approach is developed further in Hamilton et al. (2016). For all the theoretical arguments, these authors find weak empirical evidence

5. We consider 3y and 2y yields vs. short-term rates. Convexity is defined as the 5y yield vs. the average of 2y and 10y yields.

6. We refer to these indicators as real rates and real rates vs. growth.

for the link between growth and short-term real rates.

Central banks are a force to be reckoned with and modeling their decisions is part and parcel of classic macroeconomics. They target an inflation level (Bernanke and Mishkin 1997). Neglecting the impact of raw materials and other external factors, excessive inflation is directly linked to an overheating labor market. This is the Phillips curve. The relationship between unemployment and output is known as Okun's law (Ball, Leigh, and Loungani 2013). Through short-term rates, central banks hope to steer output growth, indirectly controlling unemployment and inflation.

Of course, long-term rates also play a major role in the economy and they were not directly controlled by central banks until the advent of Quantitative Easing. Bosworth (2014) investigates the determinants of 10y rates in a number of developed and emerging markets. The author warns against modeling interest rates within a closed-economy framework. His analysis points to foreign interest rates and debt/GDP as key drivers of long-term rates, with growth and unemployment not playing a consistent role.

Other authors focus on the link between short- and long-term interest rates, very often in a closed economy. So-called macro-finance models (Rudebusch 2010) aim to capture how central banks react to fluctuations in inflation, growth and possibly unemployment (Kopp and Williams 2015). These equations involve a lot of parameters, which are inferred by observing the shape of the whole interest rate curve. Ang, Piazzesi, and Wei (2006) infer growth expectations from the curve, with good explanatory power. Comparing curve-based expectations to other financial and economic measures of growth is a possible approach for the value investor.

With such a variety of approaches to the equilibrium and fair value of rates, our data set covers the most classic ones, such as time patterns in yields, 10y real rates and real rates vs. growth. We also consider CPI inflation, GDP growth and unemployment, which are key ingredients for the main macro-finance models. These fundamental indicators are sometimes compared to reference levels, typically potential growth, non-accelerating rate of unemployment, or target inflation. We consider both levels, based on the latest publication date, or changes, measured either directly or as a gap from a moving average.

All economic data are obtained from the OECD, with quarterly updates for growth[7] and monthly updates for unemployment and inflation. When simulating strategies, it is common practice to lag the data by one quarter in order to cope with revisions.

Defensive: While Asness et al. (2015) find it difficult to apply the quality concept in fixed income, Brooks, Palhares, and Richardson (2018) look to buy government bonds with the lowest duration. Leote de Carvalho et al. (2014) discuss this approach extensively. Given our investment universe, we follow a different path. The counter-cyclicality in bond returns is a well-known phenomenon, see Ludvigson and Ng (2009) for example. According to Longstaff (2004), the liquidity premium in Treasury bonds is related to a flight to quality, which occurs when consumer confidence drops, foreign investors buy more Treasury bonds and investors broadly shift funds out of equities into money markets. These observations can be justified theoretically. In the dynamic equilibrium model developed in Vayanos (2004), investors'

7. As our factors are updated every month, growth data is assumed to be constant between two updates.

preference for liquidity increases with volatility in risky assets.

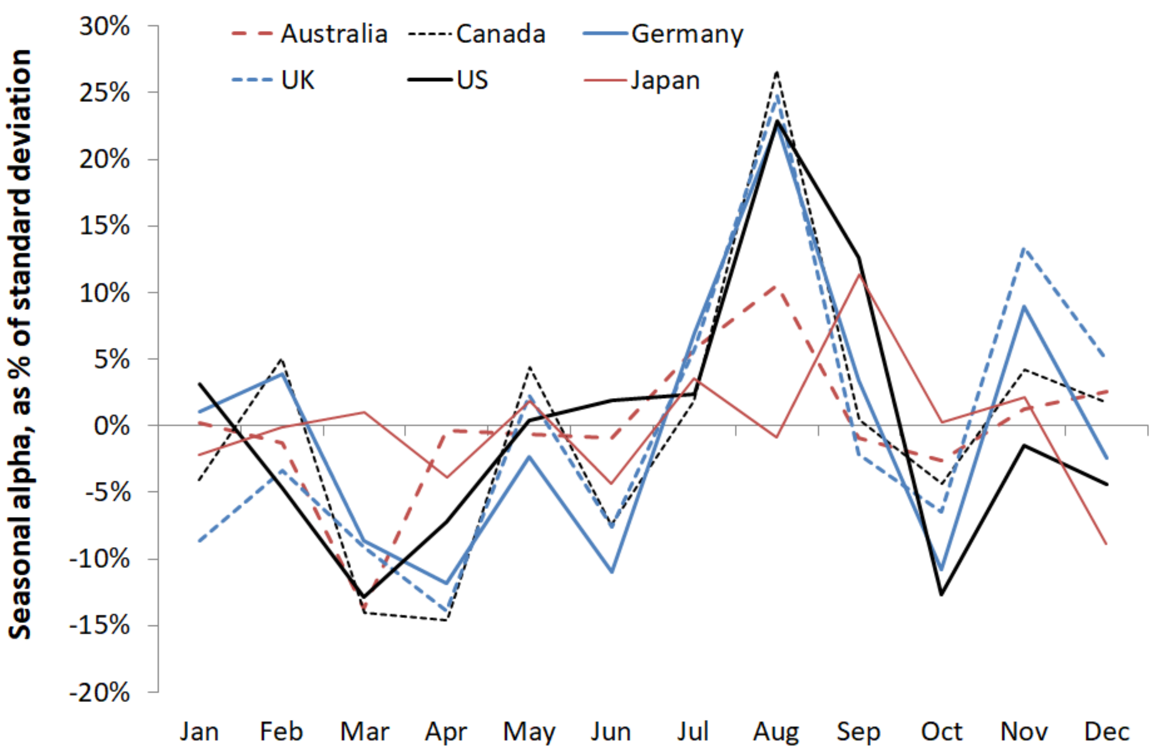
Baz, Sapra, and Ramirez (2019) show that equity market shocks are associated with flight-to-quality effects, whereas both bonds and equities fall in a bond market shock. Therefore, the stock market volatility looks like a good guide for defensive bond investors. We focus on historical volatility on local country indices, looking both at levels and changes. We do so for the Great Moderation data set only, given the lack of classic benchmark indices and the difficulty to retrieve daily data in the 1970's.

The drivers of government bond futures

Seasonality and mean reversion

Bond futures go through different cycles. It is important to identify which of these cycles may be related to alternative factors. To start with, the seasonal component in bond futures[8] is far from negligible.

Figure 1: Time-series seasonality during the Great Moderation



Source: BNP Paribas
Note: For the Inflation and Moderation data set, the patterns displayed on figure B.1 in appendix B are relatively similar, with a less pronounced peak in August.

8. We use a simple additive model with trend, based on monthly data.

Zaremba and Schabek (2017) provide an account of the academic literature related to seasonality in bond returns. They focus on two classic patterns; higher returns in January and in the six months from May to October. They find evidence of an inverted "sell-in-May" effect in 10y bonds in the US, Canada, the UK and Australia, four countries that make up a large part of our sample. These observations do not extend to other countries. They do not find evidence for seasonality playing a significant role in factor portfolios.

Schneeweis and Woolridge (1979) identify a change of pattern in the US in the early 1970's. The peak in returns shifted from January to October-November, a point they ascribe to changes in the supply and demand for credit. They cite tax laws, risk patterns and information lags as other determinants of seasonal movements. Baltas (2016) does not find relevant statistical evidence for seasonality in government bonds in a more recent data set.

Our observations point to higher returns for bonds in the second half of the year, with a peak in August. Japanese bonds do not exhibit much seasonality. Returns are usually weaker in March-April and in October. The latter effect is less pronounced in the Inflation and Moderation data set. In the cross-section, seasonal effects represent less than 5% of the alpha and tend to be corrected after 2 to 3 months. This confirms the findings of Zaremba and Schabek (2017) about the lack of seasonality in bond factors.

While seasonality in bond futures may not be a reliable source of returns, it may be hard to disentangle from mean-reversion on certain time horizons. In this part of the study, the time series of bond returns are adjusted for seasonality, while cross-sectional data is not.

This brings us to mean-reversion, which is key to some measures of value. Cumulative returns of bond futures and alphas are adjusted for trends[9]. The statistical link between variance and time horizon is a useful indication of mean-reversion[10]. We measure the half-life using the method that is described in d'Aspremont (2011).

Table 3: Half-lives of fluctuations

		Time series						**Cross section**					
Frequency	**Data***	**AU**	**CA**	**DE**	**GB**	**US**	**JP**	**AU**	**CA**	**DE**	**GB**	**US**	**JP**
Daily	GM	2.3†	1.9†	2.2†	2.2†	2†	1.9†	0.7†	1†	1.1†	1.1†	1†	1.4†
Weekly	GM	9†	8.1†	8.1†	8.1†	8.6†	5.2†	4.9†	4.4†	8.4†	4.7†	11.7†	4.7†
Monthly	IM	12.6	9.8	13.7	9.5	9.6	-	8.7	8.5	8.0	5.7	13.1	-
	GM	8.7†	7.7	9.0	8.8†	8.1	5.2	7.0	5.5	10.0	5.2	13.7	4.7
Quarterly	IM	11.7	9.4	12.2	8.1	10.3	-	10.3	8.7	8.9	5.7	12.3	-

* *GM stands for Great Moderation and IM for Inflation and Moderation.*
† *Hurst exponents stands between 0.5 and 0.2. All other exponents stand below 0.2.*
Note: The half-lives are expressed in months.
Source: BNP Paribas

9. We use a Hodrick-Prescott filter, with the adjustment proposed by Ravn and Uhlig (2002). An ADF tests confirms that the residual is stationary.

10. The Hurst index is a key indicator in the theory of self-similar processes. We use a generalized Hurst index, as described in Górski, Drożdż, and Speth (2002). A value of 0.5 or below suggests that the data reverts to the mean.

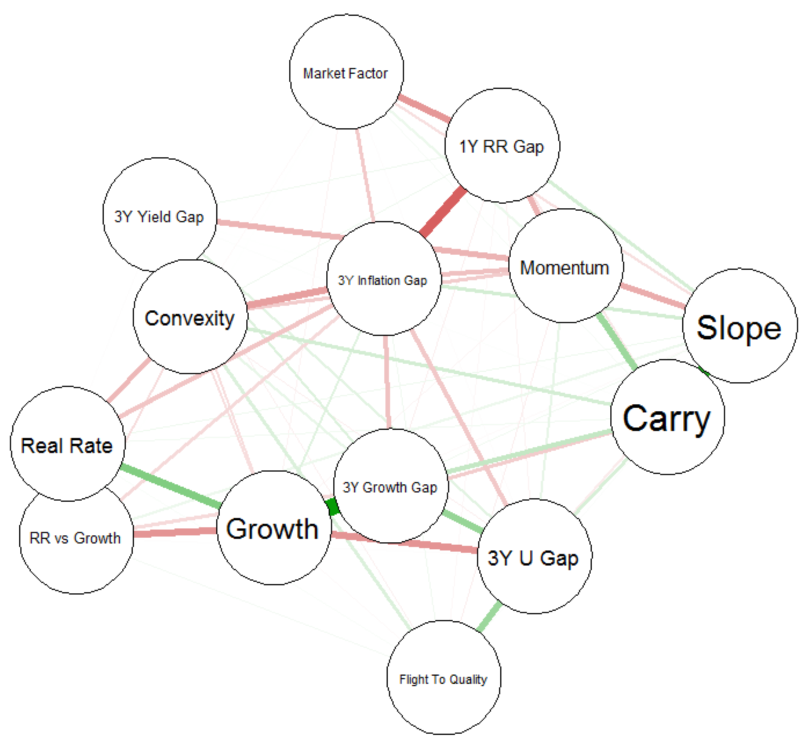
When looking for mean-reversion, the frequency of the data is key. We consider daily, weekly, monthly and quarterly returns and find evidence for stationarity in all cases. Monthly and quarterly data clearly revert to the mean, daily and weekly data a bit less so.

The fluctuations observed in daily data correct after 1 or 2 months. With monthly and quarterly returns, the horizon is about 10 months in time seris and 9 months in the cross section. In Canada, Germany and the UK, corrections occur faster in the cross-section than they do in time-series, but the opposite is observed in the US.

Selecting the right variable

Researchers usually propose a certain definition for a set of factors and then proceed to show that these factors have a large explanatory power and can be a good guide for investors. Before assessing the former point, we follow an inductive path to selecting the variables that underpin factors. Rather than measuring the explanatory power of a given set of variables, we infer from the data which combination of variables best forecasts bond returns.

Figure 2: Conditional dependencies between the main variables



Note: An edge connects two factors that are genuinely correlated conditional to the rest of the data being unchanged. The procedure, known as covariance selection, is detailed in d'Aspremont (2011). Correlation are based on quarterly returns, across all countries in the Great Moderation. RR stands for real rates, U for unemployment. Source: BNP Paribas

The table in appendix C sums up the variables that were selected for this part of the study. Economic data is not lagged, as we are looking to identify the impact of the past state of the economy on future returns. This assumption is revisited in the next part.

Figure 2 displays the most relevant connections between the main variables. One group of variables is related to the market factor and another combines the fundamental variables. Carry is connected to the former group through growth gap and to the latter through momentum.

We follow Brooks and Moskowitz (2017) in running panel regressions in time series or the cross section, using machine learning techniques for selecting variables. All returns and indicators are normalized, either across time or across countries, depending on the analysis. Coefficients are set to be the same for all countries.

Identifying leading indicators from a relatively large list is not straightforward. Brooks and Moskowitz (2017) consider 12 possible combinations of variables out of 6 indicators. Ludvigson and Ng (2009) discuss the difficulty of selecting variables in a large set of financial and economic indicators. Using principal components analysis, they identify five factors, which are related to real economic growth, interest rates, equities and inflation, in decreasing order of importance. They also find evidence of non-linearity between bond returns and the real factor. Bianchi, Büchner, and Tamoni (2019) compare various machine learning methods for predicting bond returns and find most value in neural networks with deep layers, in part due to the capacity of these techniques to capture complex non-linearities. Both studies find substantial value in macroeconomic data.

PCA analysis is a very useful technique for feature extraction and neural networks can be powerful tools for regression analysis. However, these techniques were not designed to identify individual variables. Variables selection is a very broad field of research. Subset selection requires exponentially-growing computation time and stepwise selection tends to identify sub-optimal solutions. The Lasso (Tibshirani 1996) has been widely used for selecting variables and can be conveniently solved with a fast algorithm.

The Lasso can wrongly select certain variables (Knight and Fu 2000), unless the data obeys the 'irrepresentable condition' identified by Zhao and Yu (2006) and Zou (2006). This condition also ensures that the signs of the relationships are measured correctly. However, applying it requires knowledge of the true set of variables. The adaptive Lasso (Zou 2006) has been shown to asymptotically select the right variables with the right coefficients and is much easier to use in practice.

The Lasso penalizes the fitting error by the sum of all betas, taken in absolute value. Our version of the adaptive Lasso adjusts each beta for its estimate in a simple linear regression model. This simple change allows the model to focus on the most significant variables. The penalty term is then weighted by a certain tuning parameter λ, which is selected by minimizing an adjusted Bayesian Information Criterion (Chand 2012):

$$\mathrm{BIC}_\lambda = \log(\hat{\sigma}_\lambda^2) + x_\lambda \times \frac{\log(n)}{\sqrt{n}} \tag{2}$$

where $\hat{\sigma}_\lambda^2$ is the mean square error, n is the number of dates in the sample and x_λ represents the percentage of selected variables. The resulting algorithm runs fast and is guaranteed by solid theoretical arguments to identify the right variables, provided there are enough data points. Coefficients are those of a simple linear model.

Considering quarterly data, this selection process is applied to non-overlapping returns over the next quarter, or overlapping returns over the forthcoming year (table 4). The algorithm does not retain any variable[11] for cross-sectional returns in the Inflation and Moderation data set, which is more noisy than the alternative.

Table 4: The drivers of quarterly and yearly returns

			Time series				**Cross section**	
			3m		**1y**		**3m**	**1y**
Style	**Variable**	**Sign***	GM	IM	GM	IM	GM	GM
Carry	Carry	+	+19	+20	+42	+23	+21	+11
Curve	Slope, 3y vs. ST	-	+0.3‡		+9†		-14	
Momentum	1y momentum	+				+9	+6‡	
Reversal	3y yield gap	+	+22		+47	+18		
Value	Real rate	+		+14				
	Real rate vs. growth	+				+28		
Fundamentals	3y growth gap	-				+68		-11
	Growth	-				-69		
	Unemployment	+			-27			
Defensive	2y equity vol.	+					+18	+38
R^2(%)			8	6	23	21	9	13

* *Expected sign of the beta, based on technical and fundamental patterns.*
† *p-value of the OLS regression lies between 5% and 10%.*
‡ *p-value >10%*
Note: The numbers shown are betas (%). They represent a change in annualized return as a % of standard deviation when the variable rises by one standard deviation, as measured over time or in the cross section.
Source: BNP Paribas

In line with Fattouche (2018), we find that carry matters in time series for all horizons, while value and fundamentals have a stronger impact in the longer term. We also find that momentum plays a lesser statistical role, owing to its greater variability. Reversal is more visible after one year, roughly the amount of time it takes for time series to revert to the mean (see table 3).

Similar patterns are observed in the cross-section. Momentum is a driver of quarterly returns, alongside the short end of the curve. Flight to quality, as measured by equity volatility, seems to play a key role on all horizons.

11. This happens when the marginal gain in explanatory power does not offset the additional penalty in equation (2).

The negative loading on unemployment on the left side of table 4 is contrary to economic intuition and may be related to Japan and Germany having consistently low unemployment rates.

Table 5: The top 10 drivers of bond returns

		Scores† (%)	
Style	**Indicator**	**Time series**	**Cross section**
Carry	Carry	25	27
Defensive	2y equity vol		47
Fundamental	3y growth gap	17	9
Reversal	3y yield gap	22	
Fundamental	Growth	17	
Curve	Slope, 3y vs. ST	2	12
Value	Real rate vs. growth	7	
Fundamental	Unemployment	7	
Momentum	1y momentum		5
Value	Real rate	3	

† *Percentage of total absolute betas that each variable represents in the relevant section of table 4.*
Note: Variables are ranked according to the average of the time-series and cross-sectional scores.
Source: BNP Paribas

Indicators are then ranked based on the percentage of betas they represent (table 5). Carry is key. Flight to quality is the main driver of cross-sectional returns and plays no role in time series. The short end of the curve, which reflects expectations about monetary policy, is another important source of information. Reversal and value are key factors in time series but are not selected in the cross section.

Fundamental variables play a key role in forecasting both types of returns. It is therefore interesting to take a closer look at the impact of these variables on bond returns over shorter periods of time since the early 70's. We do so for time series, as the impact of fundamentals on cross-sectional returns is somehow weaker in the longer data set. Table 6 shows that fundamental variables are not always relevant. Sometimes, as when inflation was tamed in the early 80's or during the financial crisis of 2008, classic, closed-economy variables cease to matter and other forces play a more prominent role. Classic macroeconomic relationships reasserted themselves during the tumultuous 70's and the late 90's.

Table 6: How the fundamental time-series drivers changed over time

	Start of a 5y period									
Variable	1971	1976	1981	1986	1991	1996	2001	2006	2011	2016
Growth										
3y growth gap		+22							+28	
Unemployment		-31				+45				
3y unemployment gap	+50	+54		+39						
3y inflation gap		-33				+34			+36	
R^2 (%)	25	52		15		22			25	

Note: Forecast of yearly returns, taken once every quarter over consecutive periods of 5y. A feature selection process is applied over each of these periods, focusing on our fundamental variables solely. The table displays the betas (in %) of the selected variables.
Source: BNP Paribas

A high inflation gap was a negative factor for bonds during the 70's. After the 2008 crisis, the link turned positive, as quantitative easing brought bond yields down and helped to revive inflation. It is interesting to note that the unemployment gap appears at multiple occasions in table 6 with a sign that is consistent with economic intuition.

Extracting information from fundamental variables is not straightforward. As noted by Bianchi, Büchner, and Tamoni (2019), the link between economic fundamentals and bond prices is sometimes complex and non-linear. Whether highly-dimensional statistical models are the right tools to capture this link is another question. In any case, economic fundamentals are key drivers of bond returns, alongside style and technical variables. It is now time to investigate the use of all these variables in managing cross-sectional portfolios.

Looking for evidence on factors returns

In research papers, factor portfolios are often presented without transaction costs or portfolio constraints (see table 1). Such measures are useful for identifying indicators that best discriminate across top and bottom securities. It is then necessary to estimate how much of these returns are accessible to investors, once transaction costs and management constraints have been factored in. Another question is how far such returns are linked to the market factor. A portfolio that looks nice in a study, thanks to a market bias, may not be able to withstand rising interest rates.

As we broaden the scope of available styles and factors, it is interesting to identify which ones may be genuinely relevant. Robustness across countries is a point that is sometimes overlooked. An investment style that would have worked in a diverse set of countries and macroeconomic conditions is more likely to pass the test of time. We also apply techniques that were developed by Fabozzi and López de Prado (2018) in order to cope with the selection bias under multiple testing. These techniques involve carefully identifying clusters, which are another useful piece of information in order to make sense of factors data.

In the previous part of this study, we identified a number of key variables. We now take a step back and associate a factor portfolio to each possible variable. In some cases, this is a simple choice. In others, we have to chose among a host of possible variables. We distinguish between simple and composite factors and select a representative portfolio in each case. These portfolios are our factors and are grouped into small, bottom-up clusters. A refined analysis of conditional correlations between these clusters enables us to identify four extended styles.

López de Prado and Lewis (2019) identify clusters using an unsupervised algorithm. We follow a two-step approach, selecting one factor portfolio for each composite factor first and then applying a statistical visualisation tool.

All these styles, indicators, variables, factors, clusters and extended styles are summed up in the table in appendix C.

Style investing in practice

Table 1 sums up how various related studies approached portfolio construction and transaction costs. Our long/short factor portfolios are constructed by buying and selling bond futures in proportion to a given indicator. Technically, the weight of a given country is derived from a cross-sectional score[12]. The directional bias in the resulting portfolio is removed by buying or selling the right amount of the market factor[13].

With no transaction costs or portfolio constraints, our approach amounts to allocating across alphas in proportion to the cross-sectional score of each country. Beyond transaction

12. Defined by comparing a given indicator with the cross-sectional average and adjusting the difference by the cross-sectional dispersion across indicators.

13. We look to cancel the correlation with the market factor. The time windows and frequencies as those used in the first part for measuring betas.

costs, our simulations include constraints on exposures and turn over.

With equal weight long/short portfolios, Brooks, Palhares, and Richardson (2018) report market exposures that range between 7% and 14% for value, momentum and carry. Fattouche (2018) reports lower numbers. In any case, our allocation process cancels these exposures out.

Portfolios are adjusted every time the associated variable is updated. In the Inflation and Moderation data set, data is monthly, portfolios are adjusted every month to every quarter and returns are interpolated daily. In the Great Moderation data set, data is daily and certain portfolios are adjusted every day. All measures of risks and information ratios are based on daily calculations. Compared to studies that consider monthly returns, this approach leads to technically higher estimated volatility and lower information ratios, at least for the Great Moderation data.

Table 7 presents the factors for which there was no ambiguity in selecting the right variable. One possible exception is convexity but we considered a single implementation. It is interesting to note that roll-down accounts for about half of the returns of the carry factor. Correctly measuring roll-down is likely to be key for a successful strategy.

Table 7: Simple factors

		Great Moderation				**Inflation and Moderation**			
Factor	**Sign**	**Return** (%)	**Vol** (%)	**MDD**† (%)	**IR**‡	**Return** (%)	**Vol** (%)	**MDD**† (%)	**IR**‡
Simple carry	+	+1.6	9.4	-27	+0.17	+3.8	9.8	-32	+0.39
Carry + roll-down	+	+3.4	9.6	-18	+0.35	-	-	-	-
Convexity	+	+1.6	8.0	-16	+0.20				
Growth	-	+2.4	8.2	-23	+0.29	+1.8	9.6	-40	+0.19
Unemployment	+	-2.0	7.8	-50	-0.26	-0.3	8.9	-50	-0.03
Inflation	-	+3.5	8.3	-14	+0.42	+0.5	9.9	-55	+0.05

† *Maximum draw-down*
‡ *Information ratio*
Source: BNP Paribas

Table 8 lists those indicators that involve a range of possible parameters. In each case, the table details the number of variations that were considered. For example, momentum can be measured on the returns of bond futures or those of alphas. When measuring returns, the start point is kept constant, at exactly one year from the data of calculation. However, with varying end points and smoothing periods, we end up with 16 possible variables and as many factor portfolios.

Table 8: Composite factors

Factor	$N^{\dagger}$	Great Moderation Return (%)	Great Moderation Vol (%)	Great Moderation MDD (%)	Great Moderation IR	Factor	$N^{\dagger}$	Inflation and Moderation Return (%)	Inflation and Moderation Vol (%)	Inflation and Moderation MDD (%)	Inflation and Moderation IR
Slope (-)	4										
2y vs. ST		-2.2	9.2	-53	-0.24	-		-	-	-	-
3y vs. ST		-2.6	9.2	-61	-0.29	-		-	-	-	-
7y vs. ST		-4.6	9.3	-77	-0.49	-		-	-	-	-
Momentum[‡]**(+)**	16						9				
α, 3m/ 3m		+2.5	9.7	-28	+0.26	No lag/ 1m		+0.8	10.4	-44	+0.08
3m/ 1m		+1.8	9.2	-19	+0.20	α, 3m/3m		+0.6	9.8	-38	+0.06
1m/ 1w		+0.6	9.1	-23	+0.06	α, 3m/1m		+0.2	9.8	-40	+0.02
Reversal (yield, +)	16						14				
2y gap		-0.3	8.8	-31	-0.04	3m change		+2.8	9.9	-27	+0.29
3m gap		-1.6	8.4	-51	-0.19	6m change		+1.7	10.0	-40	+0.17
1m change		-3.8	8.0	-69	-0.47	3y change		-0.1	10.5	-42	-0.01
Value (+)	2						2				
Real rate vs. growth		-0.1	8.8	-26	-0.01	RR vs. growth		+1.9	10.1	-34	+0.19
Real rate		-1.2	9.0	-45	-0.13	Real rate		+1.4	10.5	-51	+0.13
Growth gap (-)	8						8				
3y change		+2.3	8.2	-18	+0.29	3y gap		+1.6	9.8	-38	+0.16
2y gap		+1.1	8.5	-25	+0.14	1y gap		+1.0	9.9	-39	+0.10
1y change		+0.2	8.4	-25	+0.02	1y change		-0.1	9.9	-60	-0.01
Unemployment gap (+)	14						14				
9m change		+3.3	8.3	-24	+0.40	9m change		+2.7	10.1	-33	+0.26
9m gap		+2.7	8.3	-22	+0.33	3y gap		+1.3	10.1	-40	+0.13
1m change		+1.0	8.1	-26	+0.13	1m change		-0.2	8.7	-37	-0.02
Inflation gap (-)	13						13				
2y change		+1.4	9.2	-19	+0.15	2y gap		+0.9	10.5	-36	+0.08
3m gap		+0.1	8.7	-27	+0.01	3y change		-0.4	10.7	-71	-0.04
1y change		-0.4	9.2	-27	-0.05	1m change		-1.2	9.1	-65	-0.13
Flight to quality (+)	6										
3y equity vol		+4.3	7.9	-13	+0.54	-		-	-	-	-
1y equity vol		+2.8	8.2	-15	+0.34	-		-	-	-	-
1m equity vol		+0.7	8.4	-30	+0.08	-		-	-	-	-

[†] *Number of simulated strategies for a given factor. The table presents the top, median and bottom strategies based on their information ratio.*
[‡] *An α indicates that momentum is based on the alpha of futures. Start point is always 1y ago, the first number indicates the lag and the second one the smoothing window.*
Source: BNP Paribas

Table 8 also displays the top, median and bottom portfolios for each factor, based on information ratios. In order to reduce the selection bias, López de Prado (2019) considers a minimum-variance allocation across all the different versions of a strategy. As a diversified allocation may benefit from an artificially reduced level of volatility, we opt for the median portfolio. When the number of portfolios is even, we consider the higher of the two middle information ratios. In the remainder of this study, any reference to a composite factor relates to this median portfolio [14].

Our estimated information ratios are about half as high as those reported by Fattouche (2018), who also considers transaction costs. The discrepancy may be partly due to Japanese bonds not being included in our Inflation and Moderation data set. Portfolio constraints and the frequency of the data may be other factors to consider. The main discrepancy arises for value as defined by real rates vs. growth with lagged economic data. According to Brooks, Palhares, and Richardson (2018), value offers the highest information ratio. However, these authors consider forward-looking inflation expectations. Overall, all studies suggest that carry is likely to be a reliable source of returns, with more or less mixed findings for the other factors.

Table 9 sheds new light on the robustness of factors across countries. Numbers were estimated by applying the cross-sectional score of each country to its particular alpha. Without transaction costs and portfolio constraints, country-specific returns are supposed to sum up to the combined factor portfolio.

The high information ratio of carry in Japan shows that Japanese bonds can be successfully traded based on cross-sectional carry, once market exposure is hedged out. In contrast, carry does not seem to be very relevant when trading Treasuries versus other bonds.

Table 9: Information ratios across countries during the Great Moderation*

				Countries					
Style	**Factor**	**IR**	**Disp†**	**AU**	**CA**	**DE**	**GB**	**US**	**JP**
Carry	Carry	+0.35	0.18	+0.14	-0.12	-0.02	+0.27	0.00	+0.33
Curve	Slope	-0.29	0.21	-0.16	+0.04	+0.04	-0.48	-0.02	+0.09
	Convexity	+0.20	0.26	-0.00	-0.25	+0.10	+0.00	-0.12	+0.51
Momentum	Momentum	+0.20	0.17	+0.13	-0.23	+0.09	+0.09	+0.30	+0.02
Reversal	Reversal	-0.19	0.22	-0.10	-0.22	-0.10	+0.27	-0.38	+0.06
Value	Value	-0.01	0.24	+0.16	-0.07	+0.09	-0.13	+0.25	-0.42
Fundamentals	Growth	+0.29	0.15	-0.03	+0.06	+0.27	+0.05	+0.36	+0.21
	Growth gap	+0.14	0.07	+0.02	+0.11	+0.07	+0.02	+0.21	+0.03
	Unemployment	-0.26	0.31	+0.23	-0.34	-0.15	+0.04	+0.13	-0.58
	Unemployment gap	+0.33	0.17	+0.21	-0.14	+0.10	+0.05	+0.35	+0.25
	Inflation	+0.42	0.23	+0.31	-0.19	+0.24	-0.01	+0.14	+0.44
	Inflation gap	+0.01	0.09	-0.01	+0.07	+0.12	-0.09	+0.08	-0.07
Defensive	Flight to quality	+0.34	0.29	+0.16	+0.15	+0.20	-0.32	-0.19	+0.48

* *The table in appendix E displays more statistics on risks and returns across countries in the two data sets.*

† *Standard deviation of information ratios across countries*
Source: BNP Paribas

14. The reader is warned that the same factor can be represented by two different factor portfolios across the two data sets.

Dispersion, which is defined as the standard deviation of country-specific information ratios, is a useful measure of robustness. Overall, fundamental factors exhibit the lowest dispersion. Unemployment is a noticeable exception, probably because each country has its own structural level of unemployment.

In terms of robustness across countries, momentum is second to the fundamental factors only. Carry is also well ranked but the returns presented in table 9 are concentrated in 3 of the 6 countries. The table in appendix E presents more robust numbers for the Inflation and Moderation data.

Both convexity and flight to quality exhibit a high dispersion, with higher returns in Japan and Germany. In the latter case, this discrepancy may reflect structural variations in the volatility of equity indices. It may be necessary to adjust for such discrepancies in order to improve the robustness of this defensive factor.

Let us also note that fundamental factors, alongside value, are based on lagged macroeconomic data. Table 10 displays the changes in returns and risks once the data is no longer lagged. It is interesting to note that factors related to growth and unemployment are not dramatically impacted by the lag in the data. This point confirms the relevance of these factors for fixed income investors.

Table 10: Changes in returns and risks when fundamental data is no longer lagged

	Great Moderation					Inflation and Moderation			
Factor	**Return** (%)	**Vol** (%)	**MDD** (%)	**IR**	**Factor**	**Return** (%)	**Vol** (%)	**MDD** (%)	**IR**
Value									
Real rate vs. growth	+1.4	=	+6	+0.15	RR vs. growth	+0.1	+0.3	=	=
Real rate	+0.6	=	+14	+0.06	Real rate	+1.2	-6.6	+22	+0.14
Growth	+0.2	=	+9	+0.03		-2.3	+0.3	-19	-0.23
Growth gap									
3y change	-1.5	-0.2	-4	-0.19	3y gap	-1.7	=	-17	-0.17
2y gap	+0.2	-0.1	+5	+0.02	1y gap	-1.6	=	-19	-0.16
1y change	-0.5	-0.1	-9	-0.05	1y change	-1.6	+0.2	-12	-0.16
Unemployment	+0.1	=	+2	+0.02		+0.2	=	=	+0.02
Unemployment gap									
9m change	+0.5	=	+8	+0.05	9m change	+0.1	-0.2	-11	+0.02
9m gap	+0.9	+0.2	+10	+0.09	3y gap	+0.3	+0.2	-7	+0.02
1m change	+1.6	-0.1	+12	+0.20	1m change	+1.1	+0.5	+9	+0.12
Inflation	+0.6	=	+2	+0.07		+0.8	+0.2	+3	+0.08
Inflation gap									
2y change	+0.7	+0.3	+3	+0.07	2y gap	+0.5	-0.1	+1	+0.06
3m gap	+1.4	-0.3	+4	+0.17	3y change	=	=	=	=
1y change	+1.4	=	+7	+0.15	1m change	+1.4	+0.1	+17	+0.15

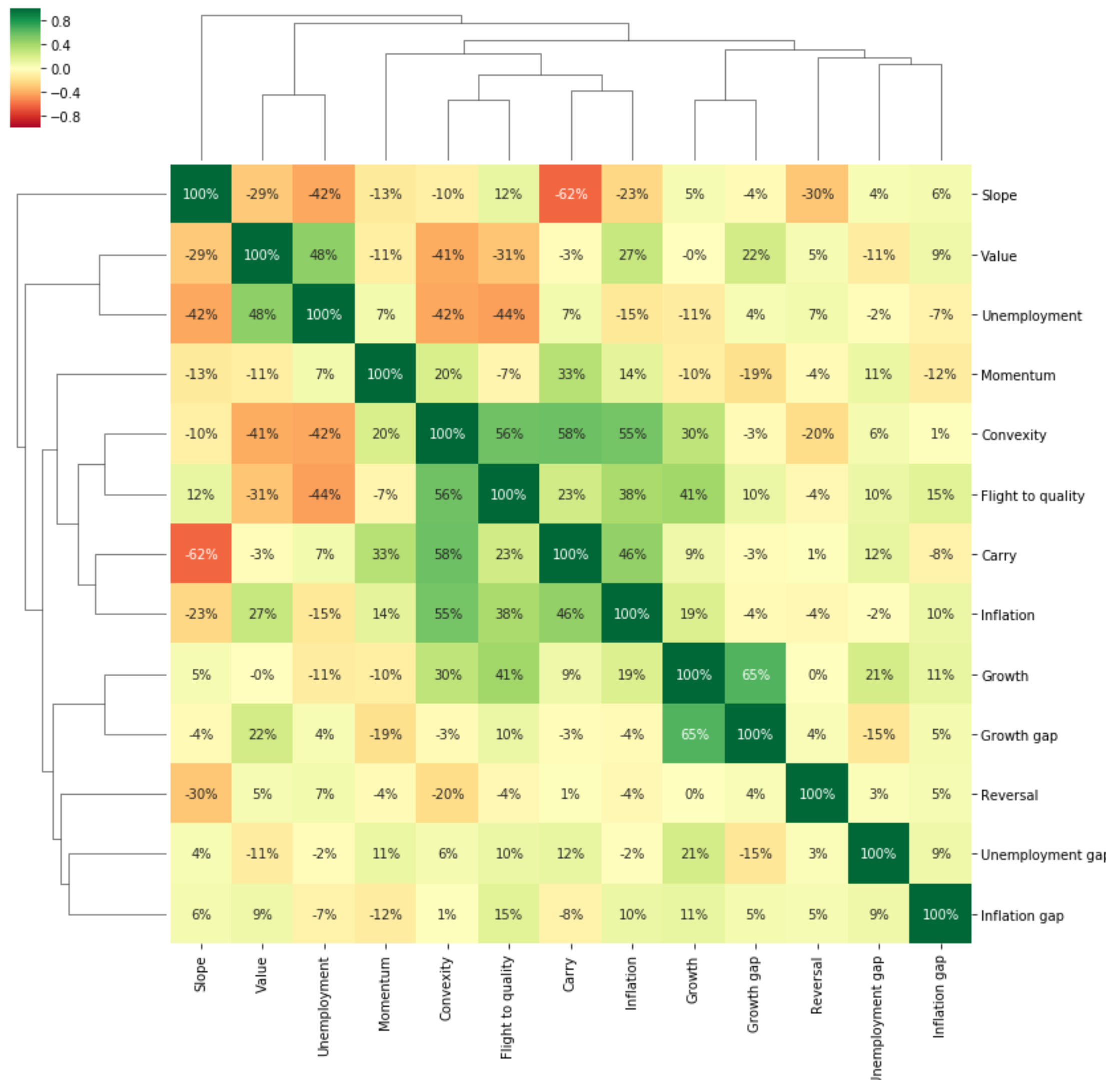
Note: The table displays changes in risks and returns when economic data is no longer lagged. A positive change means higher returns, more volatility, less negative draw-down or higher information ratio.
Source: BNP Paribas

Inflation is the factor that is most impacted by the lag. Both value factors considered in this study rely heavily on inflation numbers and would certainly be improved by a proper forecasting process. This point is visible in the higher returns reported by Brooks, Palhares, and Richardson (2018) using survey-based inflation forecasts.

Clusters and extended styles

We now look to group these factors into a set of broadly independent clusters. Figure 3 displays correlations between factors for the Great Moderation, together with a dendrogram. Four bottom-up clusters are visible. For each of them, we reiterate the same procedure as in the previous section.

Figure 3: Correlations between factors, Great Moderation



Note: Correlations of monthly returns. Source: BNP Paribas

With two factors, this leads us to select the one with the higher information ratio. Unemployment, convexity and growth gap are left aside; value, flight to quality[15] and unemployment gap are retained. Applying the same process based on figure D.1 for the Inflation and Moderation also leads to leaving aside the growth gap and unemployment. Inflation is clustered with carry and looks slightly more profitable in the Great Moderation but much less so in the Inflation and Moderation data. Therefore, carry is selected to represent this cluster.

All factors beyond slope are united within a large top-down cluster and carry appears right in the middle of figure 3. In other words, most factors benefit from a positive carry. A flat or inverted short end of the curve may be an interesting signal but following this signal can only lead to negative carry. Information stemming from the curve may be better handled through convexity, which is itself clustered with flight to quality. On this account, we leave the slope factor aside from the final list of clusters.

Table 11 sums up the correlations between the clusters that are eventually selected. The average pairwise correlation is 4% during the Great Moderation and 1% over the Inflation and Moderation periods. These numbers rise to 12% and 8% if we consider absolute values. In all cases, these are fairly low numbers. The clusters are good candidates for the procedure that is described in López de Prado (2019) and applied in the next section[16].

Table 11: Correlations across clusters

	Carry	**Momentum**	**Reversal**	**Value**	**Growth**	**Unemployment gap**	**Flight to quality**
Carry	100	+33	+1	-3	+9	+12	+23
Momentum	*+6*	100	-4	-11	-10	+11	-7
Reversal	*-16*	*-6*	100	+5	0	+3	-4
Value	*+37*	*+6*	*+9*	100	0	-11	-31
Growth	*-3*	*-8*	*+2*	*+1*	100	+21	+41
Unemployment gap	*+8*	*-7*	*-7*	*-8*	*0*	100	+10
Flight to quality	-	-	-	-	-	-	100

Note: Monthly correlations (%) during the Great Moderation appear in the upper triangular part. Correlations in the Inflation and Moderation data set are shown in the lower part. Numbers are taken from tables 3 and D.1. Source: BNP Paribas

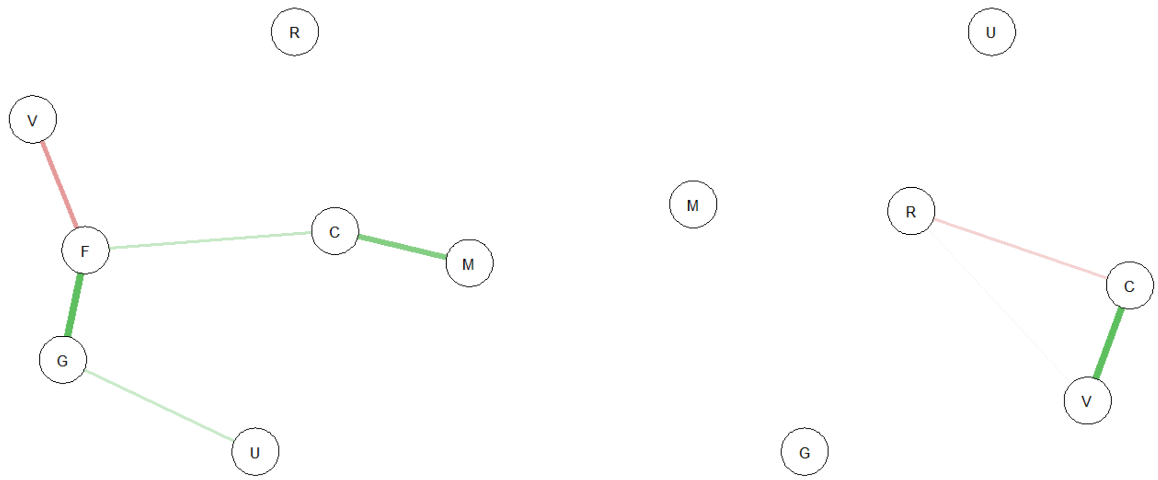
A careful look at table 11 shows that reversal is the only factor that is not positively linked with carry in any of both data sets. Value looks like a great diversifier in the Great Moderation but is highly correlated with carry in the longer term. Figure 4 sheds new light on the dependencies between the selected clusters.

15. Flight to quality and convexity are part of the same cluster. The former is linked to a higher equity volatility, the latter to a lower expected rates volatility. There is also a link between higher equity volatility and a flattening or inverting short end of the curve.

16. Bailey and López de Prado (2014) show how to adjust calculations for non-zero correlations. However, with the number of clusters and the low correlations considered here, the adjustment does not have a material impact.

Figure 4: Networks of conditional dependencies

(a) Great Moderation (b) Inflation and Moderation



C: carry, F: flight to quality, G: growth, M: momentum, R: reversal, U: unemployment.
Source: BNP Paribas
Note: The graph is based on the same covariance selection process as the one used for figure 2.

These graphs unveil the chains of conditional dependencies that underpin the correlations observed in table 11. As momentum indicators are based on past returns, carry is also a key component of the momentum factor (figure 4a). Flight to quality is also partly driven by carry, possibly for the wrong reason, namely the higher volatility in the Japanese stock market. Flight to quality is conditionally correlated with growth, itself connected with unemployment gap.

Correlations are weaker in the Inflation and Moderation dataset (figure 4b) and there is no flight to quality phenomenon to connect economic fundamentals with carry. However, the signs of correlations happen to be broadly the same in both data sets (table 11).

While the link between carry and flight to quality may be fortuitous, there seems to be a genuine connection between the latter and the two fundamental factors. After all, when growth plummets, unemployment rises and equity volatility tends to rise. All may be part of an extended defensive factor. The fundamental indicators that play a key role in the classic macro-finance models convey information on how central banks make decisions, which are often positive for bond holders in periods of stress.

In contrast, value is negatively linked with flight to quality and fundamentals, in a beautiful display of the risk-on nature of that investment style. Value seems to be decorrelated from carry during the Great Moderation but this pattern does not hold in the longer data set.

Four main groups emerge from this analysis: Carry and momentum are directly related and likely to represent the core part of a portfolio. Value is a risk-on factor that is only partially linked with carry. A broad group of factors with defensive features may help balance exposure to both carry and value. Finally, reversal is little correlated with any of the other factors, the only observable pattern being a possibly negative correlation with carry.

Which clusters are genuinely relevant?

Following López de Prado (2019), we now report skewness and excess kurtosis for the selected clusters. It is worth noting that many factors exhibit positive convexity. Kurtosis, which partly measures the risk of large movements, is much higher in the Inflation and Moderation data set.

The probabilistic Sharpe ratio (PSR)[17] represents the probability of a given Sharpe ratio being genuinely positive, adjusting for the risk of non-normal returns. The discounted Sharpe ratio (DSR) goes one step further and considers biases due to multiple testing. It is relatively easy to find high simulated returns in a large list of independent clusters. Therefore, the discounted Sharpe ratio is lower than its probabilistic counterpart. Appendix H gives more details on the calculations.

Table 12: The main clusters and their discounted Sharpe ratios

	Great Moderation						**Inflation and Moderation**					
Cluster	**IR**	**Disp**	γ_1 †	γ_2 ‡	**PSR*** (%)	**DSR§** (%)	**IR**	**Disp**	γ_1 †	γ_2 ‡	**PSR*** (%)	**DSR§** (%)
Carry	+0.35	0.18	-0.38	11	97	91	+0.39	0.06	+0.67	47	100	99
Momentum	+0.20	0.17	-0.36	15	85	69	+0.06	0.10	+0.13	44	67	50
Reversal	-0.19	0.22	-0.10	9	15	6	+0.17	0.14	+1.29	42	89	79
Value	-0.01	0.24	+0.20	9	49	28	+0.19	0.13	+1.04	57	91	82
Growth	+0.29	0.15	+0.39	11	94	84	+0.19	0.15	+0.37	34	91	83
Unemployment gap	+0.33	0.17	+0.18	9	96	89	+0.13	0.13	-1.21	52	83	70
Flight to quality	+0.34	0.29	-0.01	8	97	90	-	-	-	-	-	-

† *Skewness*
‡ *Excess kurtosis*
* *Probabilistic Sharpe Ratio associated with a rejection threshold of* 0%
§ *Discounted Sharpe Ratio*
Note: The Family-Wise Error Rate (FWER) is 92% for the Great Moderation. Reversal is an obvious culprit for this high number and the FWER drops to 65% once this cluster is excluded. The FWER is 56% for the Inflation and Moderation.
Source: BNP Paribas

17. defined as $\widehat{PSR}[0]$ using the notations of Bailey and López de Prado (2014). Skewness and kurtosis do not materially impact the calculations when information ratios are significantly lower than 1 (see equation 10).

Carry is the only factor that hits the 95% probability threshold and it does so in the longer data set. At a certainty level of about 90%, unemployment gap and flight to quality also look genuinely profitable during the Great Moderation. So, the most statistically relevant factors are carry and the extended defensive group that was identified in the previous section.

The low DSR associated with value is likely to be improved by considering forward-looking inflation expectations. Brooks, Palhares, and Richardson (2018) also note that momentum falls short of the statistical tests. However, momentum is also among the most robust factors across countries and it is worth considering this point.

The ideal factor is based on a sound economic rationale, is robust across countries and comes with a genuinely attractive estimated track record. How to combine dispersion with the discounted Sharpe ratio is an interesting topic for further research. Table 13 ranks the five most attractive factors based on a combined score.

Table 13: Looking for the right trade-off

	Great Moderation			**Inflation and Moderation**		
Rank	**Style**	**Disp**	**DSR** (%)	**Style**	**Disp**	**DSR** (%)
1	**Growth**	0.15	84	**Carry**	0.06	100
2	**Unemployment gap**	0.17	89	**Value**	0.13	82
3	**Carry**	0.18	91	**Reversal**	0.17	79
4	**Momentum**	0.17	69	**Momentum**	0.10	50
5	**Flight to quality**	0.29	90	**Unemployment gap**	0.13	70

Note: For each data set, Disp and DSR are normalised by their maximum value. We give a negative weight to the Disp score and add the two scores. Clusters are ranked by decreasing order.
Source: BNP Paribas

Momentum does not rank too bad in terms of robustness across countries. Beyond carry and momentum, value seems to have been most relevant during the 70's and 80's. For all the pitfalls in extracting information from macroeconomic fundamentals, unemployment gap appears in our top list for both data sets. Growth has been more relevant during the Great Moderation than previously. Simple reversal performs poorly over that period but looks more relevant once the 70's and 80's are factored in. Finally, flight to quality comes with a relatively high DSR but lacks robustness across countries because of structural differences between volatilities.

Conclusion

Overall, the relevant factors identified in the previous section are also those that are selected as key drivers of bond returns (table 14). Carry and the extended defensive style play a major role in both studies. The main difference lies in the relative weight of flight to quality versus fundamentals. The former plays a major role in the variable selection process, alongside the short end of the curve. The latter seem more relevant as factor portfolios.

Table 14: Relevant clusters, country robustness and statistical patterns

					Scores*	(%)
Extended style	**Cluster**	**DSR**†	**Factors**	**Disp‡**	**Cross section**	**Time series**
Carry	**Carry**	91	Carry	0.18	27	25
	Momentum		Momentum	0.17	5	
Value	**Value**		Value	0.24		7
Defensive	**Flight to quality**	90	Flight to quality	0.29	47	
			Convexity	0.26		(a)
	Growth		Growth	0.15		17
			Growth gap	0.14	9	17
	Unemployment gap	89	Unemployment gap	0.17		(b)
			Inflation gap	0.09		
Reversal	**Reversal**		Reversal	0.22	(c)	22

† *At least in the order of 90% during the Great Moderation*
‡ *Great Moderation*
* *Scores attributed to statistical drivers in table 5*
(a) *Slope obtains a score of 12% in the cross section.*
(b) *Appears as a temporary driver in table 6.*
(c) *Time pattern with an average half life of 9 months*
Source: BNP Paribas

Table 14 is also a remainder that flight to quality, based on the volatility of the stock market, is part of the same bottom-up cluster as convexity. Convexity extracts information from the short end of the curve and is theoretically connected with rates volatility. A multi-faceted approach to defensive investing would combine both technical factors with economic fundamentals. Further research is needed in order to handle country biases in the technical factors.

There is a clear case for reversal as a key time pattern, and the reversal factor happens to be relatively robust when including the 70's and 80's. However, such a simple approach struggles during the Great Moderation, as a possible consequence of central banks influencing short-term rates and subsequently long-term rates.

From both standpoints, value and reversal are of lesser importance. Even without lagging the data, value is not among the key drivers of cross-sectional bond returns. One interesting question is whether using more advanced macro economic models of the rate curve would change this observation. As fundamental factors happen to play a rather defensive role, doing so amounts to combining two practically different styles.

Combining factors and adjusting for multiple testing is another point that is worth investigating. By all accounts, carry is the most profitable factor. A factor that does not partly benefit from the carry premium is bound to underperform. It may be more relevant to assess such factors as possible tools to diversify risks in a portfolio that is centered around carry.

Endnotes

The authors would like to thank Fatma Jerad (BNP Paribas, Ecole Polytechnique) for her contribution in identifying relevant clusters in government bond factors and Marie-Stéphanie Prévot for helping edit this document.

References

Ang, A., M. Piazzesi, and M. Wei. 2006. "What does the Yield Curve Tell us about GDP Growth?" *Journal of econometrics* 131 (1-2): 359–403.

Arnott, R., C. R. Harvey, and H. Markowitz. 2019. "A Backtesting Protocol in the Era of Machine Learning." *The Journal of Financial Data Science* 1 (1): 64–74.

Asness, C. S., T. J. Moskowitz, and L. H. Pedersen. 2013. "Value and Momentum Everywhere." *The Journal of Finance* 68 (3): 929–985.

Asness, C., A. Ilmanen, R. Israel, and T. J. Moskowitz. 2015. "Investing With Style." *Journal of Investment Management* 13 (1): 27–63.

Bailey, D. H., and M. López de Prado. 2014. "The Deflated Sharpe Ratio: Correcting for Selection Bias, Backtest Overfitting, and Non-Normality." *The Journal of Portfolio Management* 40 (5): 94–107.

Ball, L. M., D. Leigh, and P. Loungani. 2013. "Okun's Law: Fit at Fifty?". National Bureau of Economic Research.

Baltas, N. 2016. "Multi-Asset Seasonality and Trend-Following Strategies." *Markets & Investors (Forthcoming).*

———. 2017. "Optimizing Cross-Asset Carry": 317–364.

Baltas, N., and R. Kosowski. 2017. "Demystifying Time-Series Momentum Strategies: Volatility Estimators, Trading Rules and Pairwise Correlations." *Trading Rules and Pairwise Correlations (May 8, 2017).*

Baz, J., N. Granger, C. R. Harvey, N. Le Roux, and S. Rattray. 2015. "Dissecting Investment Strategies in the Cross Section and Time Series." *Available at SSRN 2695101.*

Baz, J., S. Sapra, and G. Ramirez. 2019. "Stocks, Bonds, and Causality." *The Journal of Portfolio Management* 45 (4): 37–48.

Bernanke, B. S., and F. S. Mishkin. 1997. "Inflation Targeting: a New Framework for Monetary Policy?" *Journal of Economic perspectives* 11 (2): 97–116.

Bhansali, V., J. Davis, M. P. Dorsten, and G. Rennison. 2015. "Carry and Trend in Lots of Places." *The Journal of Portfolio Management* 41 (4): 82–90.

Bianchi, D., M. Büchner, and A. Tamoni. 2019. "Bond Risk Premia with Machine Learning." *USC-INET Research Paper,* no. 11.

Borio, C. E. V. 1997. *The implementation of monetary policy in industrial countries : a survey.* 148 pages. Bank for International Settlements, Monetary / Economic Dept Basle, Switzerland. ISBN: 929131045.

Bosworth, B. P. 2014. "Interest Rates and Economic Growth: Are They Related?" *Center for Retirement Research at Boston College Working Paper,* no. 8.

Brooks, J., and T. J. Moskowitz. 2017. "Yield Curve Premia." *Available at SSRN 2956411.*

Brooks, J., D. Palhares, and S. H. Richardson. 2018. "Style Investing in Fixed Income." *The Journal of Portfolio Management* 44 (4): 127–139.

Chand, S. 2012. "On Tuning Parameter Selection of Lasso-Type Methods-a Monte Carlo Study." In *Proceedings of 2012 9th International Bhurban Conference on Applied Sciences & Technology (IBCAST),* 120–129. IEEE.

Cochrane, J. H., and M. Piazzesi. 2005. "Bond Risk Premia." *American Economic Review* 95 (1): 138–160.

d'Aspremont, A. 2011. "Identifying Small Mean-Reverting Portfolios." *Quantitative Finance* 11 (3): 351–364.

Dorsten, M. P., J. Davis, and G. A. Rennison. 2016. "The Carry and Value Pendulum". PIMCO, October.

Fabozzi, F. J., and M. López de Prado. 2018. "Being Honest in Backtest reporting: A Template for Disclosing Multiple Tests." *The Journal of Portfolio Management* 45 (1): 141–147.

Fama, E. F., and K. R. French. 2015. "A Five-Factor Asset Pricing Model." *Journal of Financial Economics* 116 (1): 1–22.

Fama, E. F., and R. R. Bliss. 1987. "The Information in Long-Maturity Forward Rates." *The American Economic Review:* 680–692.

Fattouche, C. 2018. "Style Investing in Rates Markets". Barclays, February.

Górski, A. Z., S. Drożdż, and J. Speth. 2002. "Financial Multifractality and its Subtleties: An Example of DAX." *Physica A: Statistical Mechanics and its Applications* 316 (1-4): 496–510.

Hamilton, J. D., E. S. Harris, J. Hatzius, and K. D. West. 2016. "The Equilibrium Real Funds Rate: Past, Present, and Future." *IMF Economic Review* 64 (4): 660–707.

Heath, D., R. Jarrow, and A. Morton. 1992. "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation." *Econometrica: Journal of the Econometric Society:* 77–105.

Hodrick, R. J., and E. C. Prescott. 1997. "Postwar US business cycles: an empirical investigation." *Journal of Money, credit, and Banking:* 1–16.

Kessler, S., and B. Scherer. 2009. "Varying Risk Premia in International Bond Markets." *Journal of Banking & Finance* 33 (8): 1361–1375.

Knight, K., and W. Fu. 2000. "Asymptotics for Lasso-Type Estimators." *The Annals of statistics* 28 (5): 1356–1378.

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

Kopp, E., and P. D. Williams. 2015. "A Macroeconomic Approach to the Term Premium" 18/140. IMF, June.

Leote de Carvalho, R., P. Dugnolle, X. Lu, and P. Moulin. 2014. "Low-Risk Anomalies in Global Fixed Income: Evidence from Major Broad Markets." *The Journal of Fixed Income* 23 (4): 51–70.

Litterman, R., and J. Scheinkman. 1991. "Common Factors Affecting Bond Returns." *Journal of fixed income* 1 (1): 54–61.

Longstaff, F. A. 2004. "The Flight-to-Liquidity Premium in U.S. Treasury Bond Prices." *The Journal of Business* 77 (February): 511–526.

López de Prado, M. 2019. "A Data Science Solution to the Multiple-Testing Crisis in Financial Research." *The Journal of Financial Data Science:* jfds–2019.

López de Prado, M., and M. J. Lewis. 2019. "Detection of False Investment Strategies Using Unsupervised Learning Methods." *Quantitative Finance:* 1–11.

Ludvigson, S. C., and S. Ng. 2009. "Macro Factors in Bond Risk Premia." *The Review of Financial Studies* 22 (12): 5027–5067.

Ravn, M. O., and H. Uhlig. 2002. "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations." *Review of economics and statistics* 84 (2): 371–376.

Rudebusch, G. D. 2010. "Macro-Finance Models of Interest Rates and the Economy." *The Manchester School* 78:25–52.

Schneeweis, T., and J. R. Woolridge. 1979. "Capital Market Seasonality: The Case of Bond Returns." *Journal of Financial and Quantitative Analysis* 14 (5): 939–958.

Tibshirani, R. 1996. "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statistical Society: Series B (Methodological)* 58 (1): 267–288.

Vayanos, D. 2004. "Flight to Quality, Flight to Liquidity, and the Pricing of Risk". National bureau of economic research.

Zaremba, A., and T. Schabek. 2017. "Seasonality in Government Bond Returns and Factor Premia." *Research in International Business and Finance* 41:292–302.

Zhao, P., and B. Yu. 2006. "On Model Selection Consistency of Lasso." *Journal of Machine learning research* 7 (Nov): 2541–2563.

Zou, H. 2006. "The Adaptive Lasso and its Oracle Properties." *Journal of the American statistical association* 101 (476): 1418–1429.

Appendices

Appendix A Reconstructing bond futures

The "Start date" column of table A.1 gives the date from which BNP Paribas futures indices are available on Bloomberg. We used BNP Paribas futures roll indices, which include transaction costs:

- Australia: BNP Paribas bond futures AU 10Y ER index
- Canada: BNP Paribas bond futures CA 10Y ER index
- Germany: BNP Paribas EUR 10Y futures index ER
- United Kingdom: BNP Paribas bond futures UK long gilt ER index
- USA: BNP Paribas bond futures US Tsy 10Y ER index

Before these dates, the data used are the data reconstructed from the model given by equation 3.

Table A.1: Estimating the returns of government bond futures

Country	Start date	α (%)	β (%)	**Error*** (%)
Australia	01/03/2001	27	96	3.1
Canada	09/18/1989	9	68	2.1
Germany	01/02/1990	0.7	118	4.2
United Kingdom	11/19/1982	9	73	1.5
USA	05/04/1982	6	101	2.9

* *Root mean squared error between BNP Paribas index and the simulated time series, normalized by the mean of the BNP Paribas index computed on the entire time series.*
Sources: BNP Paribas, Bloomberg, Federal Reserve of St Louis, Bundesbank.

In order to extrapolate the data, we used the yields of government bonds provided by Bloomberg and interbank rates that can be found on the web pages of central banks. The excess returns of bonds are driven by a carry and roll-down component, alongside fluctuations in yields. We relate the roll-down to the curve between 10y and 3m interbank rates. Except for Australian futures, which are settled in cash, all futures allow sellers to deliver the cheapest bonds. When valuing the future, the relevant duration may not exactly be that of a 10y bond. Our model assumes that a certain percentage of the roll-down and the 10y bond duration are relevant for the return of bond futures. These two coefficients are estimated on a period of time when bond yields and futures were both available:

$$\text{daily future return} = (1 + \alpha \cdot D)(Y - r) - \beta \cdot D \cdot \Delta Y \tag{3}$$

where D and Y are the duration and the yield of a 10y bond, r is the 3m interbank rate and ΔY is the daily change in yield.

We used the yields of 10 years government bonds obtained from the Federal Reserve of St.Louis website. Those data are available from 1960 for all countries except for Australia, for which data are available from 1969. For the short rate, the 3 months interbank discount rates was used for Australia, Canada and the UK. We used the Frankfurt banks 3 months monthly average as short rate for Germany and the FEDL01 index for the US.

The complete sources of the data used are the following

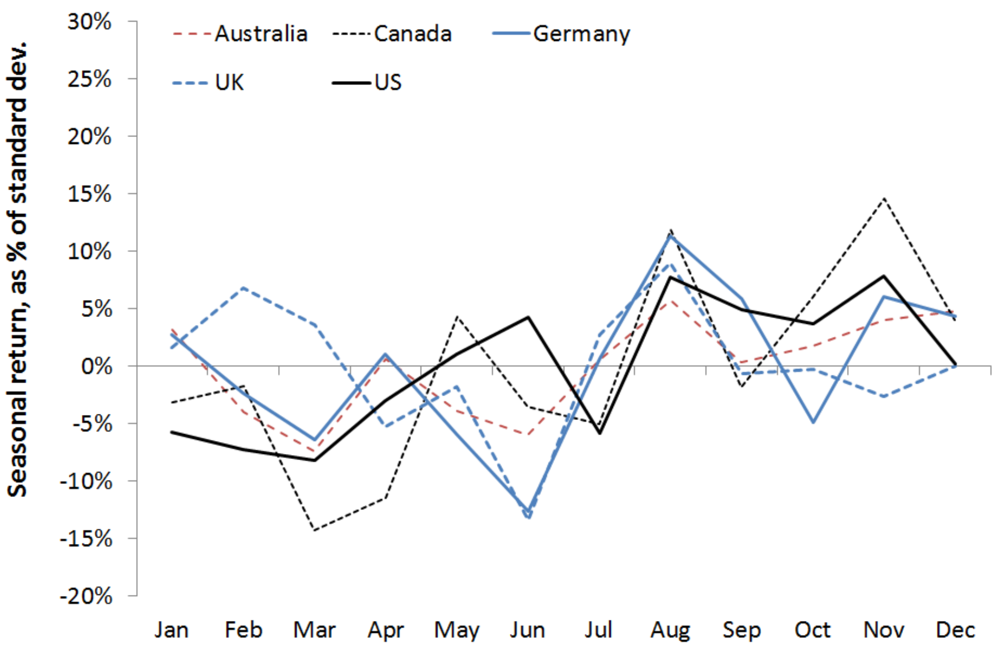
- OECD data:
  - GDP (OECD (2019), Quarterly GDP (indicator). doi: 10.1787/b86d1fc8-en (Accessed on 28 August 2019))
  - CPI (OECD (2019), Inflation (CPI) (indicator). doi: 10.1787/eee82e6e-en (Accessed on 28 August 2019))
  - HUR (OECD (2019), Harmonised unemployment rate (HUR) (indicator). doi: 10.1787/52570002-en (Accessed on 28 August 2019))
- FRED St. Louis:
  - Government yields
    - US: Organization for Economic Co-operation and Development, Long-Term Government Bond Yields: 10-year: Main (Including Benchmark) for the United States [IRLTLT01USM156N], retrieved from FRED, Federal Reserve Bank of St. Louis;
      `https://fred.stlouisfed.org/series/IRLTLT01USM156N`, August 28, 2019
    - UK: Organization for Economic Co-operation and Development, Long-Term Government Bond Yields: 10-year: Main (Including Benchmark) for the United Kingdom [IRLTLT01GBM156N], retrieved from FRED, Federal Reserve Bank of St. Louis;
      `https://fred.stlouisfed.org/series/IRLTLT01GBM156N`, August 28, 2019
    - Canada: Organization for Economic Co-operation and Development, Long-Term Government Bond Yields: 10-year: Main (Including Benchmark) for Canada [IRLTLT01CAM156N], retrieved from FRED, Federal Reserve Bank of St. Louis;
      `https://fred.stlouisfed.org/series/IRLTLT01CAM156N`, August 28, 2019

    * Australia: Organization for Economic Co-operation and Development, Long-Term Government Bond Yields: 10-year: Main (Including Benchmark) for Australia [IRLTLT01AUM156N], retrieved from FRED, Federal Reserve Bank of St. Louis;
      `https://fred.stlouisfed.org/series/IRLTLT01AUM156N`, August 28, 2019
    * Germany: Organization for Economic Co-operation and Development, Long-Term Government Bond Yields: 10-year: Main (Including Benchmark) for Germany [IRLTLT01DEM156N], retrieved from FRED, Federal Reserve Bank of St. Louis;
      `https://fred.stlouisfed.org/series/IRLTLT01DEM156N`, August 28, 2019
  - Interbank rates
    * UK: Organization for Economic Co-operation and Development, 3-Month or 90-day Rates and Yields: Interbank Rates for the United Kingdom [IR3TIB01GBM156N], retrieved from FRED, Federal Reserve Bank of St. Louis;
      `https://fred.stlouisfed.org/series/IR3TIB01GBM156N`, August 28, 2019
    * Canada: Organization for Economic Co-operation and Development, 3-Month or 90-day Rates and Yields: Interbank Rates for Canada [IR3TIB01CAM156N], retrieved from FRED, Federal Reserve Bank of St. Louis;
      `https://fred.stlouisfed.org/series/IR3TIB01CAM156N`, August 28, 2019
    * Australia: Organization for Economic Co-operation and Development, 3-Month or 90-day Rates and Yields: Interbank Rates for Australia [IR3TIB01AUM156N], retrieved from FRED, Federal Reserve Bank of St. Louis;
      `https://fred.stlouisfed.org/series/IR3TIB01AUM156N`, August 28, 2019

- Deutsche Bundesbank
  - Germany: time series BBK01.SU0107: Money market rates reported by Frankfurt banks / Three-month funds / Monthly average, retrieved from Deutsche Bundesbank;
    `https://www.bundesbank.de/dynamic/action/en/statistics/time-series-databases/time-series-databases/745582/745582?tsTab=0&tsId=BBK01.SU0107&listId=www_s510_mb03_neu&id=0`, August 28, 2019

Appendix B 50 years of seasonality

Figure B.1: Time-series seasonality during in the Inflation and Moderation data



Source: BNP Paribas

Appendix C The prospective drivers of bond futures

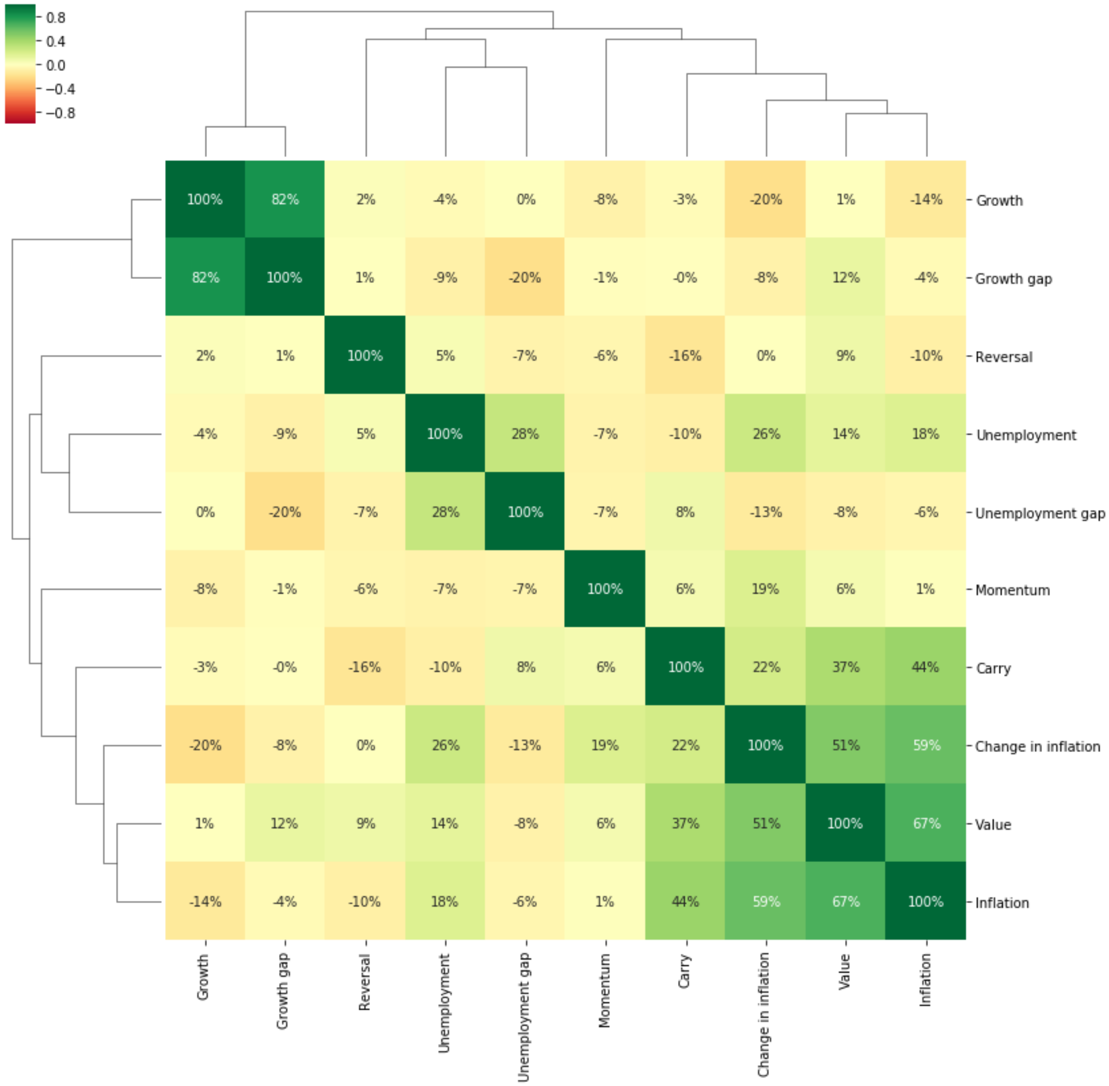
Table C.1: List of candidates for the variables selection process

		Variables for part 2†		Part 3‡		
Style	**Indicator**	**Great Moderation**	**Inflation and Moderation**	**Factor**	**Cluster**	**Extended style**
Carry	Carry	Simple carry + roll-down	Simple carry	Carry	Carry	**Carry**
Curve	Slope	Slope: 2y or 3y vs. ST, 5y vs. 2y	-	Slope*	-	-
	Convexity	Convexity (2y and 10y vs. 5y)	-	Convexity	FtQ	**Defensive**
Momentum	Momentum	1m average of returns from 1y to 1m ago		Momentum*	Momentum	**Carry**
Reversal	Yield gap	1y and 3y yield gaps		Reversal*	Reversal	**Reversal**
Value	Real rate	Real rate, level and 1y gap		Value*	Value	**Value**
	Real rate vs G	Real rate vs. growth		Value*	Value	**Value**
Fundamental	Growth	Growth, level and 3y gap		G, G gap*	Growth	**Defensive**
	Unemployment	Unemployment, level and 3y gap		U, U gap*	U gap	**Defensive**
	Inflation	3y inflation gap		I, I gap*	Carry	**Carry**
Defensive	Equity vol	Daily equity vol: 3m, 2y	-	FtQ*	FtQ	**Defensive**
		3m change in 2y vol	-			

† *"The drivers of government bond futures"*
‡ *"Looking for evidence on factors returns"*
* *Composite factors*
G: growth, U: unemployment, I: inflation, FtQ: flight to quality

Appendix D Correlations and clusters over 50 years

Figure D.1: Correlations between factors, Inflation and Moderation



Source: BNP Paribas
Note: Correlations of monthly returns

Appendix E Robustness of strategies across countries

Table E.1: Risks and returns in the two data sets

		Great Moderation				Inflation and Moderation			
Factor	Country	Return (%)	Vol (%)	MDD (%)	IR	Return (%)	Vol (%)	MDD (%)	IR
Carry	AU	+0.7	5.0	-10	+0.14	+1.3	5.0	-20	+0.26
	CA	-0.3	2.2	-11	-0.12	+0.6	2.6	-6	+0.24
	DE	-0.0	2.4	-9	-0.02	+0.8	4.2	-14	+0.18
	GB	+0.8	3.1	-6	+0.27	+0.7	3.7	-14	+0.18
	US	-0.0	3.2	-14	-0.00	+0.4	4.0	-23	+0.10
	JP	+1.0	3.1	-15	+0.33	-	-	-	-
Slope	AU	-0.8	5.0	-21	-0.16	-	-	-	-
	CA	+0.1	2.6	-8	+0.04	-	-	-	-
	DE	+0.1	2.5	-10	+0.04	-	-	-	-
	GB	-1.4	2.8	-34	-0.48	-	-	-	-
	US	-0.1	3.1	-23	-0.02	-	-	-	-
	JP	+0.3	2.9	-17	+0.09	-	-	-	-
Convexity	AU	-0.0	5.3	-22	-0.00	-	-	-	-
	CA	-0.3	1.2	-10	-0.25	-	-	-	-
	DE	+0.2	1.8	-10	+0.10	-	-	-	-
	GB	+0.0	2.0	-8	+0.00	-	-	-	-
	US	-0.2	2.0	-10	-0.12	-	-	-	-
	JP	+1.9	3.7	-8	+0.51	-	-	-	-
Momentum	AU	+0.7	5.3	-12	+0.13	+0.0	3.7	-21	+0.00
	CA	-0.5	2.3	-17	-0.23	+0.3	2.6	-13	+0.11
	DE	+0.2	2.4	-12	+0.09	-0.4	3.3	-27	-0.11
	GB	+0.3	3.0	-9	+0.09	+0.5	4.1	-27	+0.13
	US	+0.9	2.8	-7	+0.30	+0.3	5.1	-25	+0.06
	JP	+0.1	2.9	-9	+0.02	-	-	-	-
Reversal	AU	-0.5	4.5	-21	-0.10	+0.7	5.2	-22	+0.13
	CA	-0.5	2.3	-18	-0.22	+0.3	3.0	-11	+0.11
	DE	-0.2	2.2	-12	-0.10	+0.2	3.3	-22	+0.06
	GB	+0.8	2.9	-7	+0.27	+1.2	4.3	-20	+0.27
	US	-1.0	2.7	-28	-0.38	-0.5	4.0	-36	-0.12
	JP	+0.2	3.1	-9	+0.06	-	-	-	-
Real rate vs. growth	AU	+0.8	5.1	-12	+0.16	+0.9	5.0	-20	+0.18
	CA	-0.2	2.6	-12	-0.07	+0.0	2.9	-15	+0.00
	DE	+0.2	2.4	-10	+0.09	+1.0	4.0	-15	+0.26
	GB	-0.3	2.4	-15	-0.13	+0.2	3.9	-24	+0.05
	US	+0.8	3.1	-6	+0.25	-0.2	4.1	-35	-0.04
	JP	-1.3	3.1	-32	-0.42	-	-	-	-

Factor	Country	Great Moderation				Inflation and Moderation			
		Return (%)	Vol (%)	MDD (%)	IR	Return (%)	Vol (%)	MDD (%)	IR
Growth	AU	-0.1	4.8	-13	-0.03	+0.1	4.9	-27	+0.02
	CA	+0.2	2.5	-7	+0.06	+0.2	3.0	-9	+0.06
	DE	+0.7	2.5	-5	+0.27	+0.7	3.5	-13	+0.21
	GB	+0.1	2.4	-9	+0.05	+0.4	4.1	-27	+0.10
	US	+0.9	2.6	-6	+0.36	+0.4	4.0	-15	+0.11
	JP	+0.6	3.0	-8	+0.21	-	-	-	-
Growth gap	AU	+0.1	5.0	-13	+0.02	-0.1	5.0	-27	-0.02
	CA	+0.3	2.5	-6	+0.11	+0.2	2.9	-12	+0.06
	DE	+0.2	2.4	-7	+0.07	+0.4	3.6	-15	+0.10
	GB	+0.1	2.4	-9	+0.02	+0.4	4.1	-27	+0.11
	US	+0.5	2.5	-8	+0.21	+0.2	4.1	-19	+0.06
	JP	+0.1	3.0	-12	+0.03	-	-	-	-
Unemployment	AU	+0.9	4.1	-7	+0.23	+1.6	4.3	-9	+0.37
	CA	-0.9	2.7	-24	-0.34	-0.2	3.4	-24	-0.06
	DE	-0.4	2.6	-18	-0.15	-0.8	3.9	-41	-0.21
	GB	+0.1	1.8	-6	+0.04	-0.1	3.6	-23	-0.04
	US	+0.4	2.9	-12	+0.13	-0.7	4.2	-35	-0.16
	JP	-2.1	3.7	-46	-0.58	-	-	-	-
Unemployment gap	AU	+1.0	4.6	-10	+0.21	+0.7	5.0	-19	+0.15
	CA	-0.3	2.4	-17	-0.14	-0.2	2.9	-18	-0.07
	DE	+0.3	2.7	-8	+0.10	+0.7	3.6	-17	+0.19
	GB	+0.1	2.6	-7	+0.05	-0.3	4.3	-28	-0.08
	US	+1.0	2.8	-5	+0.35	+0.5	4.1	-26	+0.13
	JP	+0.7	3.0	-8	+0.25	-	-	-	-
Inflation	AU	+1.5	4.8	-14	+0.31	+0.4	5.3	-20	+0.08
	CA	-0.4	2.3	-15	-0.19	-0.3	2.7	-18	-0.12
	DE	+0.5	2.3	-8	+0.24	+1.3	4.1	-16	+0.32
	GB	-0.0	2.5	-13	-0.01	-0.9	4.1	-42	-0.21
	US	+0.4	2.7	-8	+0.14	+0.1	3.3	-24	+0.03
	JP	+1.6	3.6	-10	+0.44	-	-	-	-
Inflation gap	AU	+0.0	5.0	-23	-0.01	-0.1	6.0	-31	-0.01
	CA	-0.2	2.0	-10	+0.07	-0.2	3.0	-19	-0.08
	DE	+0.3	2.0	-8	+0.12	+0.8	3.0	-14	+0.24
	GB	-0.2	3.0	-15	-0.09	-0.1	4.0	-27	-0.03
	US	+0.2	3.0	-9	+0.08	-0.8	4.0	-53	-0.18
	JP	-0.2	3.0	-10	-0.07	-	-	-	-
Flight to quality	AU	+0.7	4.4	-8	+0.16	-	-	-	-
	CA	+0.4	2.8	-13	+0.15	-	-	-	-
	DE	+0.5	2.3	-6	+0.20	-	-	-	-
	GB	-0.4	1.4	-12	-0.32	-	-	-	-
	US	-0.4	2.1	-15	-0.19	-	-	-	-
	JP	+1.7	3.5	-8	+0.48	-	-	-	-

Source: BNP Paribas

Appendix F Trend computation using Hodrick-Prescott filter

The Hodrick-Prescott filter, defined in Hodrick and Prescott 1997 decomposes a time series $y_t,\ t = 1, 2, ..., T$ in a trend component τ_t, a cyclical component c_t and an error component ϵ_t, such that $y_t = \tau_t + c_t + \epsilon_t$. Given a properly chosen regularization parameter $\lambda > 0$, there is a trend component τ that will solve

$$\min_{\tau}\left(\sum_{t=1}^{T}(y_t-\tau_t)^2+\lambda\sum_{t=2}^{T-1}[(\tau_{t+1}-\tau_t)-(\tau_t-\tau_{t-1})]^2\right). \tag{4}$$

The first term of equation 4 penalizes the cyclical component of the signal y, while the second term is a multiple λ of the sum of the squares of the trend component's second differences. This second term penalizes variations in the growth rate of the trend component. The larger the value of λ, the higher the penalty.

In Ravn and Uhlig 2002, Ravn and Uhlig demonstrate by an empirical and analytical analysis that the Hodrick-Prescott filter parameter should be adjusted by multiplying it with the fourth power of the observation frequency ratios. This yields an HP parameter value of 6.25 for annual data, given a value of 1600 for quarterly data. The relevance of the suggestion is illustrated empirically.

Appendix G Mean-reversion half-life computation

We recall the content of the Appendix of d'Aspremont 2011, where mean-reversion estimators are computed.

In this section, we assume that we have identified an asset having a mean reverting price, and model its dynamics given by the Ornstein-Uhlenbeck process:

$$dP_t = \lambda(\bar{P} - P_t)dt + \sigma dZ_t \tag{5}$$

with P_t the price of the asset at time t, $\bar{P}$ its long-term value, σ the volatility of its price and Z_t a brownian motion modeling randomness in the price moves.
By integrating the process P_t of equation 5 over a time increment Δt we get:

$$P_t = \bar{P} + e^{-\lambda \Delta t}(P_{t-\Delta t} - \bar{P}) + \sigma \int_{t-\Delta t}^{t} e^{\lambda(s-t)} dZ_s \tag{6}$$

which means that we can estimate λ and σ by simply regressing P_t on P_{t-1} and a constant. With

$$\int_{t-\Delta t}^{t} e^{\lambda(s-t)} dZ_s \sim \sqrt{\frac{1 - e^{-2\lambda\Delta t}}{2\lambda}} \mathcal{N}(0,1), \tag{7}$$

we get the following estimators for the parameters of P_t:

$$\begin{aligned} &\hat{\mu} = \tfrac{1}{N} \textstyle\sum_{i=0}^{N} P_{t_i} \\ &\bar{\lambda} = -\tfrac{1}{\Delta t} \log\left(\frac{\sum_{i=1}^{N}(P_{t_i} - \hat{\mu})(P_{t_{i-1}} - \hat{\mu})}{\sum_{i=1}^{N}(P_{t_i} - \hat{\mu})(P_{t_i} - \hat{\mu})} \right) \end{aligned} \tag{8}$$

where Δt is the time interval between times t and $t-1$. The expression in equation 6 also allows us to compute the half-life of a market shock on P_t as:

$$\tau = \frac{\log 2}{\lambda}. \tag{9}$$

Appendix H The discounted Sharpe ratio

The following formulas are taken from Bailey and López de Prado (2014). The discounted Sharpe ratio of a strategy is given by:

$$DSR = \widehat{PSR}\left(E\left[\max\left\{\widehat{SR}_n\right\}\right]\right) = Z\left[\frac{\left(\widehat{SR} - E\left[\max\left\{\widehat{SR}_n\right\}\right]\right)\sqrt{T-1}}{\sqrt{1 - \hat{\gamma}_3\widehat{SR} + \frac{\hat{\gamma}_4 - 1}{4}\widehat{SR}^2}}\right] \tag{10}$$

where Z is the cumulative function of the standard Normal distribution, $\widehat{SR}$ is the estimated Sharpe ratio of the strategy, T is the sample length and $\hat{\gamma}_3$ and $\hat{\gamma}_4$ are respectively the skewness and the kurtosis of its returns distribution.
The expected maximum Sharpe (or information) ratio among all clusters is:

$$\begin{aligned} E\left[\max\left\{\widehat{SR}_n\right\}\right] &\approx E\left[\left\{\widehat{SR}_n\right\}\right] \\ &+ \sqrt{V\left[\left\{\widehat{SR}_n\right\}\right]}\left((1-\gamma)Z^{-1}\left[1 - \frac{1}{N}\right] + \gamma Z^{-1}\left[1 - \frac{1}{N}e^{-1}\right]\right) \end{aligned} \tag{11}$$

where $\gamma \approx 0.5772$ is the Euler-Mascheroni constant.

Under the assumption that there is no "investment skill" in the clusters, the expected Sharpe ratio $E\left[\left\{\widehat{SR}_n\right\}\right]$ is set to zero. Given that this expected value also plays a role in the definition of variance, we measured the latter as the average squared information ratio across all clusters.