Factor investing in corporate bond markets:
Enhancing efficacy through diversification and purification!

Forthcoming in The Journal of Fixed Income

Thomas Heckel
is co-head of the Quant Research Group at
BNP Paribas Asset Management in Paris, France.
thomas.heckel@bnpparibas.com, Tel. +33 (0)1 58 97 76 13

Zine Amghar
is head of Fixed Income in the Quant Research Group at
BNP Paribas Asset Management in Paris, France.
zine.amghar@bnpparibas.com, Tel. +33 (0)1 58 97 27 84

Isaac Haik
is a data scientist in the Quant Research group at
BNP Paribas Asset Management in Paris, France.
isaac.haik@bnpparibas.com, Tel. +33 (0)1 58 97 77 97

Olivier Laplénie
is head of quantitative fixed income portfolio management in the
Multi-Assets, Quant and Systematic team at BNP Paribas Asset Management in Paris, France.
olivier.laplenie@bnpparibas.com, Tel. +33 (0)1 58 97 23 31

Raul Leote de Carvalho
is deputy head of the Quant Research Group at
BNP Paribas Asset Management in Paris, France.
raul.leotedecarvalho@bnpparibas.com, Tel. +33 (0)1 58 97 21 83

August 20, 2019

Corresponding author: Raul Leote de Carvalho

Factor investing in corporate bond markets:
Enhancing efficacy through diversification and purification!

August 20, 2019

Abstract:

We show that factors from value, quality, low risk and momentum styles play an important role in explaining the cross-section of corporate bond expected returns for the U.S. and Euro Investment Grade and U.S. BB-B non-Financial High Yield universes. We demonstrate the importance of purifying factor data by neutralizing a number of risk biases that are present in the factors: controlling for sectors, option-adjusted spread (OAS), duration and size biases significantly increases the predictive power of style factors. We propose a new simple approach for efficiently neutralizing the biases from multiple risk variables and demonstrate its superiority relative to stratified sampling and optimization as alternative control methods. We also measure the added value from diversifying the number of factors in each style. Finally, we show that the results are robust in relation to transaction costs and can be used to design strategies that aim at outperforming traditional benchmark indexes.

1- Factors from value, quality, low risk and momentum styles play an important role in explaining the cross-section of corporate bond expected returns for the U.S. and Euro Investment Grade and U.S. BB-B non-Financial High Yield universes.

2- The forecasting efficacy of style factors increases significantly if biases such as sectors, OAS, duration and size in the factor data are neutralized. Diversifying the number of factors in each style also significantly improves the forecasting efficacy.

3- We propose a new simple approach for increasing the forecasting efficacy of style factors by efficiently neutralizing the biases from multiple risk variables. We demonstrate the superiority of this approach over stratified sampling and optimization.

Keywords: factor investing, smart beta, corporate bonds, credit, factor premiums, high yield, investment grade, low-risk, value, momentum, quality

JEL Classification: G11, G12, G14, E44

Factors are characteristics that play a role in explaining the returns and risk of a group of securities. Hou, Xue and Zhang (2017) listed 447 factors proposed to explain the cross-section of stock returns in the financial literature. Many of these can be grouped into styles since they are just different ways of looking at the same information. Value, quality, momentum and low risk are the styles receiving most attention probably because they are predictive of the cross-section of future stock returns and because, as mentioned by Harvey et al. (2016), the reported empirical evidence for them is more easily replicated and less likely to result from data snooping. The value style predicts higher returns for cheaper stocks, e.g. those with lower price-to-earnings ratio, as found by Basu (1971). The quality style predicts higher returns for stocks of more profitable companies, e.g. those with higher gross-profit, as discussed by Novy-Marx (2013). The momentum style predicts higher returns for stocks that have been outperforming, e.g. those with the higher 12-month average returns subtracted from the last month returns, as discussed by Jagadeesh and Titman (1993). Finally, the low risk style predicts higher returns, at least once adjusted for risk, for stocks with lower risk, e.g. those with lower historical volatility, as first found by Haugen and Heins (1972).

There is evidence that factors also play an important role in predicting the returns of fixed income securities. Haugen and Heins (1972) and later Pilotte and Sterbenz (2006) gave evidence that less risky U.S. short-dated treasury bonds deliver higher risk-adjusted returns than riskier longer-dated bonds.

Evidence that style factors also explain the cross-section of corporate bond expected returns at issue level is still limited to a relatively small number of papers written mainly by practitioners. That is surprising when the size of the corporate bond market is taken into account. According to the OECD, the total outstanding amount of corporate bonds issued by non-financial companies globally in 2018 was USD 12.95 trillion. There are a few probable reasons for this, which relate to differences between corporate bond markets and equity markets. First, at any point in time and for each company, there may be many bond issues of different maturities and with different specifications, and hence with different risk. Second, the risk of each bond changes as time passes because the time-to-maturity decreases. Third, unlike most equities, corporate bonds trade in fragmented and opaque over-the-counter (OTC) markets. Lee and Wang (2018) report a survey finding that 81% of investment grade bonds are traded by voice and Nagel (2016) report that only 5% of trading in corporate bonds takes place in all-to-all limit order books. Fourth, corporate bonds offer relatively poor liquidity because many are not easy to trade since a lot of investors buy and hold them until maturity. Fifth, the nature of debt: unlike stocks, where investors may expect to earn outsized returns from companies that see their market capitalization growing significantly over time, the returns earned by bond investors lending to those same companies is not going to be as significantly impacted by the growing equity capitalization. A company growing in market capitalization may simply take advantage of this to grow its debt by issuing more bonds without any visible impact on the returns earned by current bond holders. In first approximation, bond investors earn the yields of the bonds they buy.

The difficulty in collecting pricing data for bonds explains why databases detailing the history of prices for corporate bond markets were not available until recently. Currently, at least two datasets are available from investment banks with long traditions in bond trading: one from Bank of America Merrill Lynch (BofAML) and another from Barclays. These datasets are free of survivorship bias: whenever a company defaults, the returns of its bonds are based on their final traded price, reflecting the market's expected recovery rate. The excess returns from each corporate bond versus duration-matched Treasuries are provided. Market value, time to maturity, credit rating and credit spread are also provided.

Leote de Carvalho et al. (2014) were the first to rely on the BofAML database to provide compelling empirical evidence at bond issue level of the efficacy of low-risk factors in forecasting the cross-sectional of corporate bond risk-adjusted returns. They gave empirical evidence that corporate bonds with lower duration-times-spread (DTS) not only have an ex-post lower beta relative to the market capitalization-weighted benchmark index but also deliver higher risk-adjusted returns. Their results were observed independently for corporate bonds denominated in USD, EUR, GBP and JPY, in a study that also considered many other segments of fixed income.

Exploring factors based on fundamental data is more difficult because it requires linking the pricing data of the individual bonds of a given issuer company to the fundamental characteristics of the company in a robust manner. The fact that a particular bond keeps its name through maturity even when its issuer company changes name, e.g. because of merger or acquisition, adds to the difficulty of creating such link, in particular since relying on CUSIP and tickers is not sufficient to produce good coverage and a reliable history. In spite of this, some recent papers have already looked at the efficacy of factors that use fundamental data in explaining the cross-section of expected corporate bonds returns.

Israel et al. (2018) found that carry (option-adjusted spread (OAS)), defensive (lower leverage and higher profitability and lower duration), momentum (credit momentum and equity momentum of the bond issuer) and value factor styles (high spread relative to default risk) explain a significant portion of the cross-section of U.S. corporate bond expected returns, aggregating investment grade (IG) and high yield (HY) in one single universe. They kept only senior debt and only one bond per issuer, ignoring the issuer curve. They relied on BofAML and linked it to the Computsat fundamental company database. In order to assess the efficacy of factors, they borrowed ideas from equity factor research, e.g. Fama and French (1993), by relying on the significance of the average returns of monthly rebalanced zero-cost portfolios, long corporate bonds in the highest quintile of a given characteristic and short corporate bonds in the lowest quintile of the same characteristic. If a factor is predictive of the cross-sectional returns then the performance generated by the long-short portfolio should be positive and significant. They constructed the long-short portfolios using double sorts, on DTS first, to control for market exposure, and on the candidate factor. Finally, they showed that their results were robust to transaction costs and can be used to create investable portfolios.

Houweling and Zundert (2017) also investigated how style factors explain the cross-section of corporate bonds expected returns. Like Israel et al. (2018), they considered an approach based on zero-cost long-short portfolios, but they used deciles instead of quintiles. Unlike Israel et al. (2018), they did not control for market exposure, they used Barclays instead of BofAML, they kept all constituents of the Barclays indices, and they investigated U.S. IG and U.S. HY separately. They considered size (total market value of all bonds issue by a company), value (percentage difference between the actual credit spread and the fitted fair credit spread calculated from a cross-sectional regression on rating dummies, time to maturity, and the three-month spread change for each bond), low risk (shorter-dated and higher rated) and momentum (past-six-month return calculated with one-month implementation lag based on excess return versus duration-matched Treasuries). Their analysis did not include factors that require fundamental company data. For IG, they found that long-short portfolios formed using low risk and value factors delivered positive and statistically significant performances. They did not find momentum and size effects in IG. For HY, performance of value and momentum was positive and highly significant and for low risk was positive and significant. Finally, like Israel et al. (2018), they also concluded their results were robust to transaction costs and could be used to create investable portfolios.

In this paper we add to the literature in a number of ways. First, we consider separately the three largest universes in the BofAML database, U.S. IG, Euro IG and U.S. B-BB non-Financial HY. Performing the analysis on three separate universes allows us to test for the robustness of the results while limiting data snooping. We also diversify significantly more than previous studies in terms of the number of factors used in each style. We consider 33 factors across the four styles - value, quality, low risk and momentum - with a number of those constructed with fundamental data from the Worldscope database and analyst forecasts of fundamental data from the IBES database. Factor diversification was defended in a recent equity factor study by Leote de Carvalho et al. (2017) as a means to reduce the risk of data snooping. We believe this is even more important in credit markets since the history of the bond price datasets is shorter than for equities.

Second, we pay significantly more attention to the neutralization of unwanted risk exposures than Houweling and Zundert (2017) and Israel et al. (2018). In particular, we neutralize sector biases in factor data by de-meaning the factor data, a process common in equity factor studies that comprises subtracting the sector mean of the factor in the cross-section of a sector from all the bonds in that sector. It is puzzling to us that previous studies did not control for sector exposures. Value, quality and low risk factor data often comes with sector biases because of structural differences in sectors. But companies in different sectors have different exposures to macroeconomic shocks which impact returns independently of this factor data. Thus, neutralizing sectors should increase the efficacy of style factors since these cannot be expected to predict shocks to macro variables. This argument was put forward and illustrated by Leote de Carvalho et al. (2018) in their equity factor research. We also neutralize both duration and OAS as a means of reducing market exposure, showing that controlling for DTS is less efficient. Finally, we control for size exposures. Neutralizing size has been standard in equity factor research since the work of Fama and French (1993), who proposed neutralizing size exposures in their value factor.

Third, we propose a new and simple approach to increase the efficacy of factors in predicting the cross-section of corporate bond returns by allowing for the efficient control of multiple risk variables such as duration, OAS and size. We call it the *local scoring* approach because it is based on calculating factor scores in a small sub-universe of bonds that have similar exposures to the control variables. This approach can be seen as an improvement over stratified sampling. While in stratified sampling the sub-universes are hard-defined and each includes a pre-defined number of bonds and bounds for the control variables, in the local scoring approach the sub-universe is bond specific and defined by making sure the bonds accepted in its peer sub-universe do not deviate too far from the values of the control variables. We demonstrate the increased efficacy of this approach relative to both stratified sampling and to the use of optimization for controlling exposures to risk variables. The use of the local scoring approach also makes it easy to include the entire issuer curve.

This paper is organized as follows. We start with a discussion of data, which is one of the key challenges for credit factor investing. In particular, we discuss the question of linking fundamental data with bond price and bond characteristics. This is followed by a proposal and motivation of the choice of the factors we find pertinent for predicting the cross-section of corporate bond returns. We then describe the methodology used in our research, starting with the filters that are applied to the database to remove bonds that cannot be compared with the majority of bonds available because of their different seniority or liquidity. We describe how to standardize factor data into scores that are comparable across factors, and how to purify factors, i.e. how to neutralize unwanted risk biases that may exist in the factor data, such as sectors, OAS, duration and size biases. We propose three approaches for neutralizing OAS, duration and size biases: i) stratified sampling, ii) a new simple approach we call local scoring that can be seen as an improvement over stratified sampling, and iii) optimization. We also discuss why it is important to control for such risk biases and why failing to do so may lead to disappointing results. The last section is a discussion of results. We compare the three approaches used to neutralize risk variables and find that the local scoring method is superior to the other two. We thus use this approach in the reminder of the paper. We find that relying on a diversifying set of factors adds significant value over using only a few selected factors. We also find that, on average, factors from the four factor styles play an important role in explaining the cross-section of corporate bonds expected returns with value and momentum factors exhibiting higher efficacy. Finally, we provide an example of how to transform factor scores into long-only portfolios that are capable of outperforming traditional benchmark indexes even when taking into account the typical high transaction costs for trading corporate bonds.

DATA

All specific data for bonds such as prices, duration, maturity, OAS and ratings were taken from the BofAML database. As in other studies, the returns of bonds were considered in excess of the government bond returns of similar duration taken directly from the BofAML database, reflecting only the credit component of each corporate bond and not their interest-rate component. The BofAML database offers extensive coverage of a complete range of individual bond issues across all liquid bond markets and is widely used by fixed income asset managers for benchmarking purposes and the calculation of portfolios' net asset values.

We used indices from ICE BofAML and relied on their history of index constituents. Factor performances were measured for three largest credit universes with bonds of comparable seniority: Merrill Lynch U.S. Investment Grade index (C0A0), Merrill Lynch Euro Investment Grade index (ER00) and Merrill Lynch U.S. BB-B non-financial High Yield index (H4NF). For U.S. High Yield, we excluded bonds from financial companies, since most of them are subordinate debt, and bonds rated CCC or lower, since the issuers are too close to default and thus their risk is highly idiosyncratic and dependent on recovery rates. For this reason, we used the Merrill Lynch index H4NF instead of H0A0. Due to the small number of issuers and issues in the Euro HY universe, we did not include it in our analysis.

For the IG universes we used data starting in January 2000 whereas for the HY universe we chose January 2003 as the starting point due to the small number of issues and issuers in this universe prior to that date. We used monthly data and for all three universes the data runs through December 2017.

Fundamental data for the company issuers was taken from Worldscope database and analysts' forecasts of fundamental data was taken from the IBES database. Both databases were linked to the BofAML database to ensure the link of each bond to its obligor through time, which required carefully taking into account the impact of mergers and acquisitions throughout history. This was of key importance to ensure the correct assessment of the relevance of factors based on fundamental company data behind a given bond. A proprietary methodology for linking the different databases was used for this purpose. It required cross-referencing information also in Bloomberg, FactSet and, in a number of cases, going back to the prospectus of the corporate bond. Size data was also taken from the Worldscope database. We used the sum of all interest bearing and capitalized lease obligations of an issuer, i.e. the sum of long and short-term debt.

CREDIT FACTORS

We classified the factors into four different factor styles: value, quality, low risk and momentum. Such factors are known to explain a significant portion of the cross-section of stock expected returns. Since the fundamental value of bonds and equities both depend on the value of the underlying assets of the firm (e.g. Merton (1974)) it makes sense to expect such factor styles to explain some of the cross-section of corporate bond expected returns, too. However, because there are important differences between stock and bond prices we chose factors that, at least in our view, are more likely to be pertinent for corporate bonds.

One important difference between stocks and bonds is that while equity investors may earn significantly large returns from investing in stocks of companies that successfully grow their businesses, this kind of upside is generally not possible in the corporate bond market where many such companies will just increase the size of debt by issuing more bonds. As a first approximation, bond investors earn the yield of the bond for as long as the underlying company does not default. It is thus intuitive that corporate bond investors should focus on avoiding companies that may default. For this reason, we often use the size of the company's debt in the factor definitions.

Unlike the studies of Houweling and Zundert (2017) and Israel et al. (2018), we prefer to use more factors for each style even if they are capturing the same effect. Leote de Carvalho et al. (2017) showed that using more factors than less for each style, even when they are relatively similar, reduces the risk of data snooping and increases the overall efficacy of equity factor models. We believe the same should apply to the corporate bond market.

The value style is about investing in cheap securities. We considered two types of value factors. The first favors bonds with a larger OAS relative to a fair value OAS obtained from a cross-sectional regression at a given date of the log (OAS) against the time to maturity of each bond, and the distance to default (D2D) of each company. D2D is estimated in the manner proposed by Chuan and Wang (2012) as the number of standard deviations the asset value is away from default. Bonds with an OAS larger than fair value OAS are more attractive as their OAS is expected to close the gap. The second type of value factor aims at identifying value traps, i.e. companies that may appear to be cheap based on traditional value measures such as book-to-price, cash-flow yield, earnings yield or sales yield, but are in fact cheap because they are closer to default. Avoiding the companies with the larger values of these ratios will likely exclude such companies. While for equity investors this may sound counter intuitive, remember that bond investors do not capture an upside: they just earn the yield for as long as there is no default. Thus, between two bonds with the same spread and duration, we prefer the bond of the company with a higher price being paid by equity investors for its book value, earnings or cash-flow because equity investors are pricing a lower probability of default by showing confidence in the future of the company.

For the quality style we also consider two types of indicators. The first measures the ability of a company to cover its debt with generated income. We prefer companies that generate higher cash-flow, free cash-flow, gross profit and EBITDA relative to the size of their debt. Similarly we prefer companies with higher EBITDA relative to their interest expenses. We also prefer bonds from companies with lower accruals (calculated by subtracting either operating cash-flow or free cash-flow from net income) relative to assets. A second type of quality indicator avoids aggressive businesses. We avoid companies that are increasing either their expenditure or debt relative to their assets, or that have higher amounts of financing cash relative to their debt, or that are increasing their assets at the fastest pace. Factors based on cash-flow were not used for financial companies as a large component of their businesses is generated by interest payments not taken into account in the cash-flow.

The low risk style is about earning higher risk-adjusted returns from lower risk companies. We use two types of risk measures. The first is related to the bond market. We prefer bonds of companies with lower leverage and with higher distance to default as defined by Duan and Wang (2012). The second is related to the risk of the company in the equity market. We prefer bonds from companies with lower beta and lower historical volatility of stock returns.

Exhibit 1: list of credit factors classified into styles

Factor style	Description	Factor	Prefer high/low factor values
Value	Prefer cheaper firms	Spread relative to distance to default	high
	Avoid value traps	Book to price	low
		Cash-flow to price	low
		Reported earnings to price	low
		IBES earnings forecast to price	low
		Sales to price	low
Quality	Prefer firms capable of covering debt with generated income	Cash-flow to debt	high
		Free cash-flow to debt	high
		Gross profit to debt	high
		EBITDA to debt	high
		EBITDA to interest expenses	high
		Accruals (Op) to total assets	low
		Accruals (Fr) to total assets	low
	Avoid aggressive issuers	Capital expenditures relative to total assets	low
		Change of debt relative to total assets	low
		Financing cash to debt	low
		Annual percentage change in total assets	low
Low risk	Prefer less indebted firms	Leverage	low
		Distance to default	high
	Prefer less risky firms	Stock beta (historical 3-year weekly returns)	low
		Stock volatility (historical 3-year weekly returns)	low
Momentum	Prefer firms with stronger medium-term equity momentum and weaker short-term equity momentum	12 months – 1 month momentum	high
		12 months – 1 month alpha	high
		12 months – 1 month information ratio	high
		12 months – 1 month Jensen information ratio	high
		6 months – 1 month momentum	high
		Momentum relative to the 52 weeks high	high
		1 month reversal momentum	low
	Prefer firms with stronger fundamental momentum	6 months momentum in earnings revision	high
		12 months momentum in earnings revision	high
		Annual change in standardized IBES long term earnings growth forecast	high
		Annual change in standardized earnings	high
		Annual change in standardized free cash-flow	high

For the momentum style, we use two types of momentum. The first is based on the performance of the company in the equity markets. Much like for stocks, we prefer bonds from companies that outperform their peers in the equity markets in the medium term (six to 12-month horizon). But in the short term, we prefer bonds from companies that just underperformed, as it is the case for equities. Jensen information ratio is the average of the residuals of a regression of stock excess returns against market capitalization index excess returns divided by the volatility of those same residuals. The second type is fundamental momentum. We prefer bonds from companies with stronger earnings revisions momentum, a stronger annual change in long-term earnings growth forecasts standardized by the annual volatility of those same forecasts, a stronger annual change in earnings per share standardized by the volatility of the annual changes in earnings per share, and an annual change in free cash-flow standardized by the volatility of the annual changes in free cash-flow.

METHODOLOGY

To test the efficacy of the credit factors in Exhibit 1 in predicting the cross-section of corporate bonds returns, we relied on monthly rebalanced long-short portfolio strategies where the portfolios are long the bonds with higher expected returns and short the bonds with lower expected returns. Positive and significant average returns for such portfolio strategies indicate that the factors do appear to have predictive power for the cross-section of returns. In the first step we shall not include transaction costs as we are simply interested in assessing the efficacy of the factors in forecasting returns. It is only at the end that we construct a more realistic long-only strategy with turnover constraints and assess the impact of realistic transaction costs to investigate whether the results can be exploited from a practical point of view.

Universe

We first define the universe of bonds in which the strategy can invest. For that, we start with all bonds in the respective Merrill Lynch index at a given date and remove all non-senior debt because these bonds are not comparable to senior bonds. Then we remove low face value bonds for liquidity reasons and bonds with too long a time to maturity because there are too few of them. We also remove bonds from issuers for which we do not have information about the total size of their debt because many of the factors in Exhibit 1 use debt in their definition. Finally, we remove bonds that we consider to be outliers. We do so by comparing each bond to its issuer credit curve: if the OAS is too far from the curve then we remove the bond as this is likely due to liquidity issues, e.g. stale price of the bond, or a problem with the link between the bond and its issuer which may be wrong.

The filtered universe for U.S. Investment Grade is composed of about 1,000 bonds from 350 issuers at the start of the period and about 3,000 bonds from 700 issuers at the end of 2017. For Euro Investment Grade there are about 250 bonds from 150 issuers at the start of the period and about 1,000 bonds from 350 issuers at the end. The U.S. BB-B non-Financial High Yield universe includes about 200 bonds from 100 issuers at the start and 600 bonds from 300 issuers at the end.

Factor scores

For each corporate bond we calculate a cross-sectional score using a z-score transformation of a given factor f:

$$z_i = \frac{f_i - \langle f_i \rangle}{\sigma_f} \tag{1}$$

where, at a given date, the score z_i is calculated from the cross-sectional values, f_i, their mean, $\langle f_i \rangle$, and their cross-sectional standard deviation σ_f. The scores are standardized so that we can combine them when constructing multi-factor linear combinations. The scores are calculated so that bonds with higher scores have higher expected returns.

Neutralization of sector biases

Bonds in the same sector are more likely to be impacted similarly by shocks from macroeconomic variables that are not related to the factors we are considering here. Controlling for sectors reduces the noise generated by those risk variables and leads to more robust testing of the efficacy of the credit style factors in Exhibit 1. Thus, we neutralize sector biases in factor data by de-meaning the factor data. This is done by subtracting the sector mean of the factor in the cross-section of a sector from all the bonds in that sector.

Neutralization of OAS, duration and size biases

It is possible that the long-short portfolios constructed from factors exhibit strong OAS and duration biases. These are well-known risk variables that also impact corporate bond returns. Thus, it is important to neutralize exposures to OAS and duration if we wish to assess the efficacy of the style factors beyond any duration or OAS biases they may create. That is why we also control explicitly for OAS and duration. Ben Dor et al. (2007) found that controlling for duration times spread (DTS) should be sufficient to control for market exposure from a risk point of view. We compare the efficacy of controlling just for DTS or controlling for both OAS and duration. We find that the latter is superior as a form of controlling for the impact of these risk variables. Finally, we also control for the size, which is a variable that could impact liquidity and thus returns.

We considered three approaches that allow for the control of multiple variable biases. The first two are relatively simple: stratified sampling and *local scoring*. In these, the scores are calculated for sub-samples of the investment universe created in such way that only bonds with similar values of the control variables are kept in each sub-universe. By construction, the scores will necessarily show relatively small biases to the control variables. The latter, local scoring, is a new approach we propose as a generalization of stratified sampling to: i) render it more efficient in neutralizing the biases to the control variables; and ii) allow for the control of a larger number of variables. For both these two approaches, the weights of bonds in the long-short portfolio are simply proportional to the scores. The proportionality constant is the same for all bonds and through time. The third approach is based on using optimization to neutralize biases by brute force. Here, the scores are calculated across the entire investment universe (global scores) and will keep biases to the control variables found in the factors. It is the optimizer that imposes the neutralization of the biases to the control variables in the step of portfolio construction. We compare the three approaches.

Stratified sampling consists of partitioning the universe to form sub-universes using the variables to control and then applying the scoring methodology to each sub-universe independently. When there is only one variable to control, the approach often goes under the name of double sorts, where the universe is first ranked by the variable to control and partitioned into sub-universes and then ranked and scored by the factor in question. Biases such as OAS, duration or size can be controlled individually using this double sort approach. When more than one variable are controlled, the universe is ranked by all variables and then segmented into sub-universes by defining break points for each variable. Unfortunately, as the number of variables to control increases, the size of the sub-universes decreases and can quickly become too small for statistical analysis. The scores are calculated from equation (1) and applied to each sub-universe. We set the weight of each bond in the long-short portfolio equal to the bond score.

Local scoring is an alternative approach we propose as an improvement over stratified sampling. In this approach, a different peer sub-universe is defined for each individual bond by keeping only bonds with values of the control variables sufficiently close to those of the bond in question. This can be done by calculating the distance of a given bond to all other bonds in the universe and retaining in the peer sub-universe only other bonds for which the distance is sufficiently small. For the control variables $v_1, v_2, \ldots$ we can define the distance of a bond to other bonds by:

$$d\,(b, b') = \frac{\left|v_1^b - v_1^{b'}\right|}{v_1^b} + \frac{\left|v_2^b - v_2^{b'}\right|}{v_2^b} + \ldots \qquad (2)$$

Then, for each bond b we select a pre-defined number or percentage of bonds in its vicinity, i.e. with the shortest distances. We can subsequently apply the calculation of the score following equation (1) to all bonds selected for the peer sub-universe and retain the score of the bond in question. The process is repeated for all bonds in the universe in order to calculate a score for each bond. This approach can be used to control for OAS, duration and size simultaneously. As with stratified sampling, we set the weight of each bond in the long-short portfolio to equal the bond score.

Global scores and optimization can be used as an alternative approach to calculate the scores and construct the long-short portfolio while controlling for biases. Here, the scores are calculated from equation (1) applied to the entire investment universe. The scores are allowed to keep their OAS, duration and size biases. Then, the optimization program is asked to create the long-short portfolio by maximizing the product of each bond weight in the portfolio with its respective factor score z_i:

$$w_i = \text{argmax}(\textstyle\sum_i w_i z_i) \qquad (3)$$

under constraints $\|w_i\|_2^2 = 1$. It can be shown that the solution to this optimization problem is a long-short portfolio where the weight of each bond is proportional to its score. OAS, duration and size can be controlled by adding additional constraints explicitly, namely:

$$\sum w_i OAS_i = 0 \qquad (4)$$

$$\sum w_i Duration_i = 0 \qquad (5)$$

$$\sum w_i Size_i = 0 \qquad (6)$$

A long-only portfolio can also be constructed from this optimization program by changing the portfolio constraints. In the exercise we run at the end of next section, we use scores calculated from the local scoring approach in the optimization and impose the following criteria: a long-only constraint; that the portfolio is fully invested; that the portfolio's weighted OAS, duration and size is within 5% difference of the OAS, the duration and the size of the equally-weighted investable universe, respectively; that the weight of any bond issue must be lower or equal to 1% and the weight of any issuer lower or equal to 3%; and finally, we include an aversion to turnover aimed at keeping annual turnover in the range of 70% to 100% one way, taking into account the natural turnover of each index which is not negligible.

RESULTS

We shall now discuss the results from historical simulations carried out for the three corporate bond universes: U.S. Investment Grade, Euro Investment Grade and U.S. BB-B non-Financial High Yield. The first objective is to assess the efficacy of the credit factors in Exhibit 1 in forecasting the cross-section of bonds returns in each of these universes by applying the different long-short portfolio strategies described in the methodology section. We first compare the three approaches proposed in order to choose the one best adapted for the exercise. All information ratios presented in tables are based on the average annual alpha generated from the long-short portfolio strategies divided by the annualized volatility of the residual of the regression used to calculate the alpha. In this regression, we use the monthly returns generated in historical simulations by the portfolio strategies against the returns of the equally weighted universe. The returns used for each corporate bond to generate the results are all in excess of duration-matched Treasuries.

Stratified sampling versus local scoring

We started by comparing stratified sampling with the local scoring approach. Each of the approaches is run individually for each factor. For stratified sampling, we partitioned each variable into quintiles, forming 25 sub-portfolios of different OAS and duration levels. For the local scoring, we keep the nearest 2.5% of the total number of bonds that lie in the vicinity of each bond when applying the approach to the U.S. IG universe, and 5% for Euro IG and for U.S. BB-B non-Financial HY universe.

In Exhibit 2, we show the average of the information ratio estimated for the 33 factors in Exhibit 1. We also show the average of the absolute beta of the long-short factor portfolios over time and over the 33 factors. We control for OAS and duration only, and not size, because the stratified sampling partitions would be too small to accommodate the three variables.

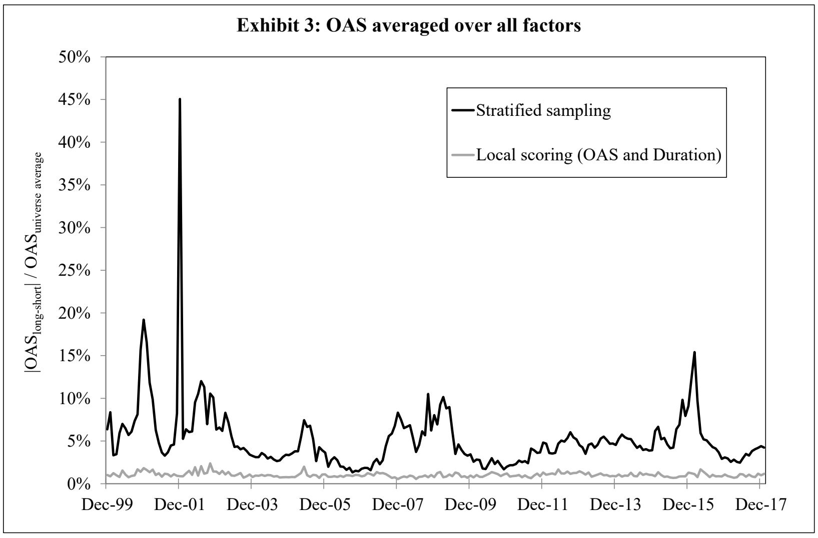
Exhibit 2: Information ratio and beta averaged over all factors; sector, OAS and duration neutral

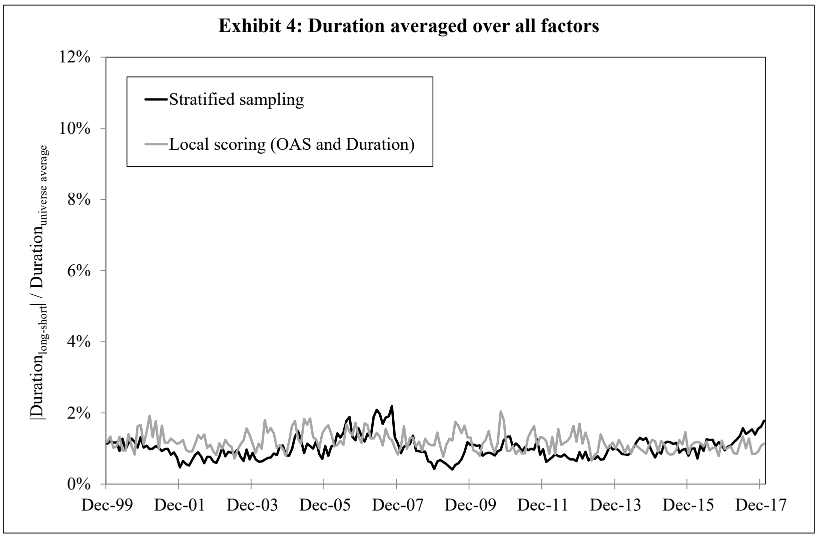
		U.S. IG	Euro IG	U.S. HY*
IR	**Stratified sampling**	1.11	0.69	0.61
	Local scoring	1.24	0.86	0.71
|Beta|	**Stratified sampling**	0.12	0.17	0.05
	Local scoring	0.06	0.07	0.04

*U.S. BB-B non-Financial HY

The first important remark is that the average efficacy of the credit factors in Exhibit 1 is strong, in particular for U.S IG, as demonstrated by the information ratios. The efficacy for the smaller universes Euro IG and U.S. HY is comparable, stronger for Euro IG when using the local scoring approach. The second remark is that the information ratios obtained from the local scoring approach are higher everywhere. Finally, controlling for OAS and duration is seen as a way of controlling for market exposure. We find that the absolute value of the beta is closer to zero when using the local scoring approach. All this, plus the limitations of the stratified sampling for many variables, plays in favor of the local scoring approach.

In Exhibit 3 we compare the efficacy of stratified sampling against the local scoring approach in controlling for OAS. We plot the absolute OAS of the long-short factor portfolios averaged over all factor portfolios generated from the 33 factors, respectively, divided by the universe average OAS, at each date. The results are for U.S. IG. In Exhibit 4 we calculate the equivalent measure for duration. As shown, the local scoring performs much better at controlling for OAS than stratified sampling. On several occasions the average absolute OAS of the long-short factor portfolios deviates by more than 5% relative to the average OAS of the universe. When it comes to controlling for duration both approaches do well, producing an absolute duration of less than 2% of the average duration of the universe. Results for the other two universes are not shown but were comparable.





Optimization versus local scoring

We focus now on the comparison of local scoring with the optimization approach described in the methodology section. Because of the flexibility of both approaches we can control for OAS, duration and size. The results are show in Exhibit 5.

Exhibit 5: Information ratio and beta averaged over all factors; sector, OAS, duration and size neutral				
		U.S. IG	**Euro IG**	**U.S. HY***
IR	**Optimization**	1.09	0.76	0.70
	Local scoring	1.27	0.92	0.71
|Beta|	**Optimization**	0.11	0.09	0.06
	Local scoring	0.07	0.07	0.04

*U.S. BB-B non-Financial HY

We first observe that the additional control of size on top of OAS and duration increases the information ratios derived from the local scoring when compared to Exhibit 2, even if only marginally. The second observation is that the information ratios for IG are significantly higher when applying local scoring than when relying on optimization. Finally, we find that the control of the risk variables leads to an absolute beta closer to zero when relying on local scoring.

The key difference between the use of global scores in the optimization approach and the use of the local scoring approach is that, by construction, the local scores are more likely than global scores to represent the entire universe. Indeed, the fact that we compare each bond only to bonds in its close vicinity should more likely lead to a larger dispersion of bond scores across the entire range of the control variables. Using global scores in optimization is more likely to meet constraints by brute force while not necessarily representing well the entire investment universe. The simplicity of the local scoring approach relative to the use of an optimizer and the superior results obtained from it motivate its choice for the remainder of this paper.

DTS versus OAS and duration neutralization

Before going any further, it is important to clarify whether market exposures are better neutralized by controlling for the DTS or explicitly by controlling for OAS and duration simultaneously. In Exhibit 6, we show the results of the local scoring approach implemented in these two ways. The results show higher information ratios for the simultaneous control of OAS and duration than by controlling only for DTS, i.e. assuming that from a risk point of view a bond with high duration and low OAS is comparable to a bond with low duration and high OAS provided that duration times OAS is the same. We also find that, on average for the factors considered, the absolute beta is slightly closer to zero when controlling for the two variables separately. For this reason, we shall continue to control for OAS and duration separately in the remainder of the paper.

Exhibit 6: Information ratio and beta averaged over all factors, sector neutral

		U.S. IG	Euro IG	U.S. HY*
IR	DTS neutral	0.95	0.77	0.52
	OAS and duration neutral	1.21	0.87	0.68
|Beta|	DTS neutral	0.10	0.12	0.04
	OAS and duration neutral	0.08	0.09	0.04

*U.S. BB-B non-Financial HY

Impact of neutralization of biases

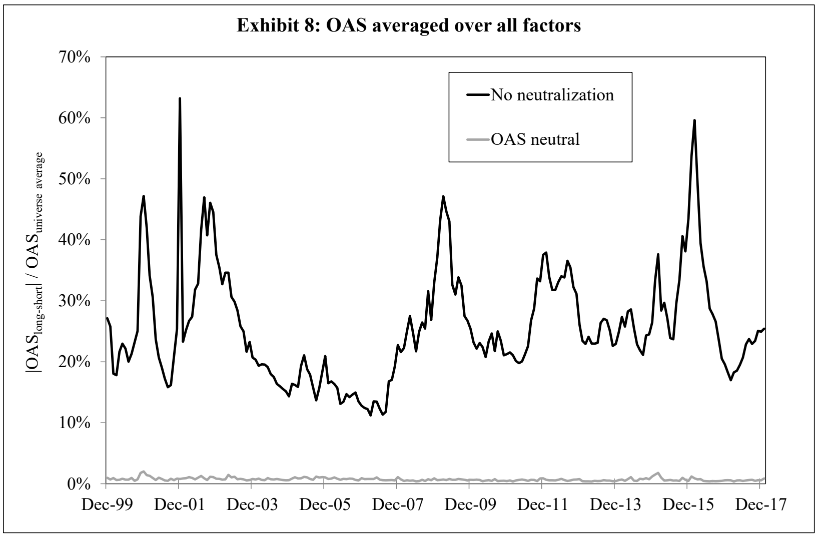
We now focus the importance of neutralizing each individual bias. We rely on the local scoring approach and first neutralize one bias at a time. The sector neutralization is achieved through de-meaning – as explained in the methodology section – while the other variables are controlled by comparing each bond to those in their vicinity only when calculating the factor scores. In Exhibit 7, we show the average information ratios arising from applying the methodology to each individual factor in Exhibit 1.

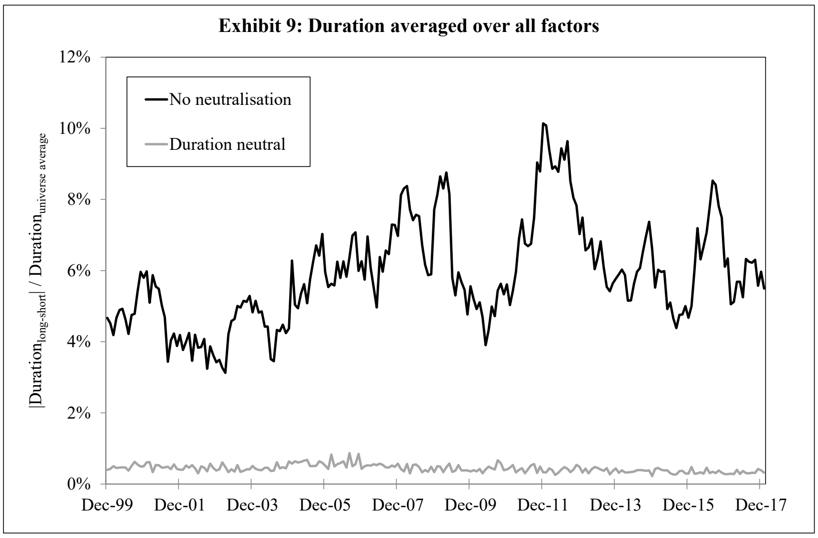
Exhibit 7: Information ratio averaged over all factors

	U.S. IG	Euro IG	U.S. HY*
No neutralization	0.63	0.54	0.35
Sector neutral	0.69	0.56	0.37
Duration neutral	0.70	0.53	0.35
OAS neutral	1.04	0.75	0.66
Size neutral	0.63	0.61	0.38

*U.S. BB-B non-Financial HY

The results show that the neutralization of biases either improves or adds no value. Controlling for OAS brings the most benefit by increasing information ratios rather significantly in all universes. Controlling for sectors, duration or size brings occasional benefits.





In Exhibit 8 we plot the absolute OAS of the long-short portfolio strategy divided by the universe average OAS when no variable is controlled and when OAS is controlled. Results are for U.S. IG. In Exhibit 9 we plot the equivalent measures for the duration, when duration is controlled. What we find is that the factor data introduces significant OAS and durations biases if neutralization is not performed. We also find that the local scoring approach, despite its simplicity, does remarkably well at controlling the biases of these variables, which is reassuring. Results for the other two universes were comparable.

Exhibit 10: Information ratio averaged over all factors

	U.S. IG	Euro IG	U.S. HY*
No Neutralization	0.63	0.54	0.35
Sector neutral	0.69	0.56	0.37
Sector, OAS neutral	1.10	0.81	0.70
Sector, OAS, size neutral	1.14	0.88	0.75
Sector, OAS, duration neutral	1.21	0.87	0.68
Sector, OAS, duration, size neutral	1.27	0.92	0.71

*U.S. BB-B non-Financial HY

We now look at the incremental value of neutralizing each risk variable. In Exhibit 10, we show the average information ratio of all factors as we increase the number of control variables. More often than not, the information ratios increase as we go down the table. If we compare the last row with the first one in Exhibit 10, we find the information ratios almost doubling in value. This is evidence that the neutralization of the biases significantly increases the average efficacy of the factors in Exhibit 1 in predicting the cross-section of bond returns.

Multi factor portfolio

So far, we have focused on the average efficacy of the factors in Exhibit 1. An important question is what the joint efficacy would be if we were to combine the factors to create a multi-factor model? Here, the long-short portfolio is constructed from the average of the factor scores calculated for each bond using the local scoring approach applied to each individual factor. When calculating scores for a given bond, we first equally weight the factor scores z_i from each factor in each factor style to produce a combined style score, and then we equally weight these aggregate style scores to form the final multi-factor score. We also consider the approach without neutralizing control variables and compare it to the approach where all risk variables are controlled. The results are show in Exhibit 11.

Exhibit 11: Information ratio				
		U.S. IG	**Euro IG**	**U.S. HY***
No neutralization	Average over the 33 factors	0.63	0.54	0.35
	Multi-factor (family weighted)	1.18	1.13	0.68
Sector, OAS, duration and size neutral	Average over the 33 factors	1.27	0.92	0.71
	Multi-factor (family weighted)	2.35	2.17	1.85

*U.S. BB-B non-Financial HY

The efficacy of the multi-factor model is rather high. It is remarkable that an exercise that was carried out by importing ideas from the equity markets, adapting them to the corporate markets using intuitive arguments, and imposing a sensible neutralization of risk variables leads to such significant results.

Factor styles

So far, we have looked at the average of efficacy from the 33 factors in Exhibit 1 and their efficacy when combined into a multi-factor model. One question not yet addressed is how the different factor styles contribute to the strong efficacy of the multi-factor model. In Exhibit 12, we show the information ratios obtained by equally weighting each factor in each style while using the local scoring approach. The results show that all factor styles have predictive power with momentum being the strongest followed by value. Quality and low risk also show significant predictive power but smaller than for momentum and value. That means that all factor styles contribute positively to the robustness of a multi-factor model.

Exhibit 12: Information ratio: sector, OAS, duration and size neutral			
	U.S. IG	**Euro IG**	**U.S. HY***
Value	2.11	1.77	1.48
Quality	1.33	0.96	0.96
Low Risk	1.74	1.34	0.99
Momentum	2.37	2.14	1.57
All	2.35	2.17	1.85

*U.S. BB-B non-Financial HY

From empirical evidence to practice

After concluding that the factors in Exhibit 1 would have predicted the cross-section of corporate bond returns in the three universes investigated, the question now is how these factors could have been used to create realistic investment strategies that outperform benchmark indexes even when taking into account transaction costs. To answer this question, we run a simple exercise where constrained long-only fully invested portfolios are constructed every month using the local scoring to form a multi-factor score as described in the previous section and an optimizer to construct the long-only portfolio and manage constraints as described in the methodology section.

In Exhibit 13, we show the results from those simulations, gross and net of transaction costs. The transaction costs for trading each bond (bid-ask spread) are given by $0.05 \times \text{OAS} \times$ duration, with OAS floored at 0.2% and capped at 6%. For a 2-year duration and OAS equal to 0.5%, the cost of one transaction is 0.05%. For a duration of 10 years and OAS equal to or larger than 6%, the cost of one transaction is 3%.

Exhibit 13: performance and risk of long-only portfolio strategies built from factor scores

	U.S. IG		Euro IG		U.S. HY*	
	Portfolio	Benchmark	Portfolio	Benchmark	Portfolio	Benchmark
GROSS						
Annualized excess return over cash	6.01%	4.12%	3.50%	2.97%	7.73%	6.37%
Sharpe Ratio	1.06	0.77	1.10	0.88	0.98	0.80
Alpha	1.88%		0.86%		1.59%	
Information ratio	1.03		0.80		0.93	
NET						
Annualized excess return over cash	5.52%	3.95%	3.24%	2.85%	6.71%	5.79%
Sharpe Ratio	0.97	0.74	1.02	0.85	0.86	0.73
Alpha	1.56%		0.71%		1.14%	
Information ratio	0.86		0.66		0.66	
Beta	1.00	1.00	0.89	1.00	0.96	1.00
Volatility	5.69%	5.36%	3.17%	3.35%	7.83%	7.90%
Tracking error	1.82%		1.07%		1.73%	
Turnover (one way)	73%	29%	73%	33%	107%	56%

*U.S. BB-B non-Financial HY

The results in Exhibit 13 show that, with a relatively simple optimization algorithm used to build a monthly rebalanced long-only portfolio from factor scores and imposing a sensible control of turnover, we can create a strategy capable of outperforming the market capitalization weighted benchmark indexes, even when taking into account realistic measures of transaction costs. This demonstrates that it should be possible to rely on the factors' efficacy in predicting the cross-section of corporate bond returns to construct implementable investment strategies designed to outperform market capitalization indexes.

CONCLUSIONS

In this paper, we proposed a list of 33 factors from value, quality, low risk and momentum styles and showed that they play an important role in explaining the cross-section of future corporate bonds. We find that diversifying the number of factors used in each factor style increases their efficacy and reduces the risk of data snooping. To our knowledge, we are the first to give such strong evidence for such a large number of factors and for three different universes: U.S. IG, Euro IG and U.S. BB-B non-Financial HY.

We propose a new and simple approach – 'local scoring' – to standardize factors so that biases for multiple risk variables can be efficiently neutralized. This approach is based on calculating the score for each bond by taking into account only other bonds with values of the control variable that are sufficiently close to that of the bond in question. We show that two alternative approaches, stratified sampling and optimization, deliver inferior results to those from local scoring when used as means to neutralize biases.

Furthermore, we emphasize the importance of controlling for risk variables that are known also to play a role in explaining corporate bond returns as a means to increase the efficacy of the factor models. In particular, we propose that sectors, OAS, duration and size exposures are neutralized and demonstrate the superiority of the results when style factors are purified in this way.

Finally, we show that our results are robust to realistic transaction costs and can be used to construct strategies designed to outperform traditional benchmark indexes.

DISCLAIMER

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

REFERENCES

Basu, S. 1977. "Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis." *Journal of Finance* 32 (3): 663-682.

Ben Dor, A., L. Dynkin, J. Hyman, P. Houweling, E. van Leeuwen, and O. Penninga. 2007. "DTS^{SM} (Duration Times Spread)." *The Journal of Portfolio Management* 33 (2): 77-100.

Fama, E. F., and R. F. Kenneth. 1993. "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics* 33 (1): 3–56.

Harvey, C. R., L. Yan, and H. Zhu. 2016. "… and the Cross-Section of Expected Returns." *The Review of Financial Studies* 29 (1): 5-68.

Haugen, R.A., and A.J. Heins. 1972. "On the Evidence Supporting the Existence of Risk Premiums in the Capital Markets." Working Paper, http://ssrn.com/abstract=1783797.

Hou, K., C. Xue, and L. Zhang. 2017. "Replicating Anomalies." NBER Working Papers 23394, National Bureau of Economic Research, Inc., http://www.nber.org/2018LTAM/hou.pdf.

Houweling, P., and J. van Zundert. 2017. "Factor Investing in the Corporate Bond Market." *Financial Analysts Journal* 73 (2): 100-115.

Israel, R., D. Palhares, and S. Richardson. 2018. "Common Factors in Corporate Bond Returns." *Journal of Investment Management* 16 (2): 17-46.

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

Duan, J.-C., and T. Wang. 2012. "Measuring Distance-to-Default for Financial and Non-Financial Firms." *Global Credit Review* 2: 95-108.

Lee, T., and C. Wang. 2018. "Why trade over-the-counter? When investors want price Discrimination." Working paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3087647.

Leote de Carvalho, R. L., 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.

Leote de Carvalho, R. L., X. Lu, F. Soupe, and P. Dugnolle. 2017. "Diversify and Purify Factor Premiums in Equity Markets." In *Factor Investing, From Traditional to Alternative Risk Premia*, edited by E. Jurczenko, 73-97. ISTE Press Ltd.: Elsevier Ltd.

Merton, R. 1974. “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates.”*Journal of Finance* 29 (2): 449-470.

Nagel, J. 2016. “Markets committee electronic trading in fixed income markets.” Working paper, Tech. rep. 9789291974207, Bank for International Settlements.

Novy-Marx, R. 2013. “The Other Side of Value: The Gross Profitability Premium.” *Journal of Financial Economics* 108 (1): 1-28.

Pilotte, E.A., and F.P. Sterbenz. 2006. “Sharpe and Treynor Ratios on Treasury Bonds.” *The Journal of Business* 79 (1): 149-180.