University of Zurich UZH

Department of Banking and Finance
Center of Competence for Sustainable Finance

ESG and Alternative Factors in Corporate Bond Returns

Master Thesis in Banking and Finance

Tobias McCarthy

Full Text Version

CCSF Thesis Series no. 25 (2021)

May 2021

CCSF Thesis Series (formerly CSP/CMF Thesis Series)

ESG and Alternative Factors in Corporate Bond Returns
Master Thesis in Banking and Finance

Tobias McCarthy
Advisor: Dr. Annette Krauss
Professor: Professor Dr. Marc Chesney

Full Text Version

CCSF Thesis Series no. 25 (2021)

Zurich: University of Zurich, Department for Banking and Finance / Center of Competence for Sustainable Finance, Plattenstrasse 14, 8032 Zurich, Switzerland

Table of Contents

List of Tables

1. Introduction

From 2000 to 2018, the total number of sustainable assets under management expanded from under $2tn to just over $30tn globally.[1] As of the beginning of 2020, sustainable assets in the U.S. alone reached $17.1tn, which represents 36% of professionally managed assets and a 42% increase since the end of 2018 (USSIF, 2020). Over the same time period, net new money flows to ESG ETFs and mutual funds (those that consider environmental, social and corporate governance criteria) increased six-fold (USSIF, 2020).[2] In conjunction with growing public awareness, regulatory pressure, and consumer demand, a shift in investor attitudes towards sustainability underpins these remarkable trends.[3] Recent surveys suggest that money managers are increasingly adopting sustainability strategies voluntarily (Andrews and Vunaki, 2021).[4] Amel-Zadeh and Serafeim (2018) find that a growing percentage of senior investment professionals from conventional funds (i.e. non SRI funds) perceive ESG integration as more of a means to enhance investment performance than as an objective to balance against financial returns.[5] While the upward trend in ESG integration is undeniable, empirical studies reveal a gap between investors' stated intentions and the degree to which investors implement ESG strategies in practice.[6] Surveys from the CFA Institute, HSBC Asset Management, and Amel-Zadeh and Serafeim (2018) indicate that investors' lack of knowledge regarding the best ways to incorporate ESG partially accounts for this divide.

The increase in the number of academic studies related to ESG parallels the dramatic growth in sustainable assets under management. According to Cappelle-Blancard and Monjon (2012), the number of publications related to corporate social responsibility (CSR) remained at a low level for decades before spiking in the early 2000s. For example, the number of articles published in 2009 (~650) exceeded the number of articles published between 1982 to 2002. Moreover, Friede et al. (2015) find that, of the 2,200 empirical studies they review, 90% report a non-negative relationship between ESG and market and/or account-based performance. These results, however, belie a distinct lack of consensus concerning the connection between ESG criteria and corporate financial performance (CFP) (Gerard, 2018). Researchers attribute these disagreements to different procedures for

---

[1] USSIF 2020 defines sustainable assets as those that investors manage with sustainability-related investment strategies.

[2] Examples of ESG factors include climate change and greenhouse gas emissions (Environmental), community outreach and labor conditions (Social), and corruption, executive pay, and board composition (Governance).

[3] European governments have implemented significant ESG reform in the past few years. The IORP II Directive (UE 2016/2341) is one such example. It establishes specific measures that govern ESG-related disclosures for pension fund beneficiaries and provides incentives fund managers to integrate ESG factors within investment policies and analyses. In contrast, the SEC in the U.S. recently passed rulings that will limit the ability for ERISA-governed pension plans to utilize ESG criteria (USSIF, 2021). Overall, however, the trend in global regulatory reform has been supportive of ESG (Seltzer, 2020).

[4] In previous years institutional investors cited non-pecuniary motivations and client demand as the primary factors behind their adoption of ESG strategies (USSIF, 2014). McKinsey finds that, compared to ten years ago, twice as many investment professionals (83%) expect the impact of ESG programs on shareholder value to increase over the next five years. https://www.mckinsey.com/business-functions/sustainability/our-insights/the-esg-premium-new-perspectives-on-value-and-performance,

[5] Specifically, respondents to the authors' survey cite the financial materiality of ESG as the primary reason they consider ESG information in investment performance (Amel-Zadeh and Serafeim (2018)).

[6] CFA Institute. 2015. Environmental, Social and Governance (ESG) Survey. Investors reveal gap between ESG belief and action | The Asset. Although now outdated, Vyvyan (2007) find that investors rate financial returns as less important in funds with strong environmental profiles, but that their actual investment preferences imply that profit maximization remains their primary concern.

data collection and analysis, inconsistent measurements of CSP and CFP, and variations in industrial and geographic contexts (Brammer and Millington, 2012, Bektic, 2018, Drempetic et al., 2019, Liu and Taylor, 2019, Berg et al., 2020). These issues are even more severe for non-equity asset classes, which have not benefitted equally from the wave of professional and academic interest in ESG (Gerard, 2018).

Research into ESG integration in fixed income has gathered steam in the past couple years, but after decades of neglect, the gap between equities and bonds is not closed easily.[7] Although bonds represent 32% of sustainable investments, Friede et al. (2015) identify only 36 studies (out of ~2200) that examine the relationship between ESG criteria and bond yields and/or returns.[8] Moreover, what little evidence there is concerning bonds and their issuer's ESG score is mixed (see Bektic, 2018). The UNPRI identifies a similar divide among its signatories, with only 19% claiming to explicitly consider the ESG profile of their corporate bond portfolios, compared to 80% with respect to equity portfolios (UNPRI, 2016).[9] These numbers are particularly shocking in the context of the economic and financial significance of the corporate bond market. As Gourio (2013) explains, for many companies it is the bond market, not the equity market, which represents the "marginal source of finance".[10] This is even truer today, as structurally lower interest rates have fueled a massive expansion in both corporate borrowing and in demand for corporate issues from institutional and retail investors (IMF, 2019, Anand et al., 2020). As a result, at the same time as the corporate bond market nearly doubled in size from 2008 to 2018 (from \$5.1tn to \$9.2tn), so too did the share of bond mutual funds and ETFs (from 9.3% to 22.5% of the total corporate bond market).[11]

The disparity in ESG-related research between equities and bonds is consistent with the general lack of research on the time-series and cross-sectional determinants of corporate bond returns (see Houweling and van Zundert, 2017 or Bai et al., 2018). Factor-based investing, which has redefined investment management in the equity market (Ang et al., 2009), remains in a nascent state in fixed income (Slimane et al., 2020). As Israel et al. (2018) explain, while researchers can choose from among well-established models in equities (e.g. the dividend discount model), there is no consensus about how to measure expected bond returns. This is partially attributable to comparatively opaque pricing, limited access to data, and a general lack of transparency in the asset class (Soe and Xia, 2016). Even after the introduction of TRACE data in 2002, progress has been sluggish, and to assess the performance of bond funds and portfolios, researchers rely on a combination of factors constructed from bond indices (i.e. *DEF* and *TERM* from Fama and French (1993)), as well as those from the equity market (Bai et al., 2018). To evaluate synthetic corporate bond portfolios, however, the former are relatively ineffective (Bali et al., 2019). This is especially the case with samples, like ours, that separate IG and HY bonds, and which exclude treasury bonds and bonds with option features (i.e. structured notes, mortgage-backed securities, and puttable bonds). Regarding the latter, rational pricing factors, or those based systematic risk premia and behavioral biases, should be consistent across equity and

[7] The majority of research into ESG and corporate bond returns has been published in the past five years (Henke, 2016, Leite and Cortez, 2016, Pereira, 2019, Hoepner and Nilsson, 2020).
[8] This still represents a step up from earlier periods. Of the 52 and 103 studies reviewed by Ortlizky et al. (2003) and Margolis and Walsh (2003) respectively, not one examines the link between CSR and the cost of debt.
[9] Slimane et al. (2019) attribute this partially to the different objectives of bond investors, which they claim consider fixed income the field of impact investing. They point to the high demand for green and social bonds to support this assertion.
[10] See Goldstein et al. (2017) for additional commentary on the importance of the corporate bond market as a source of corporate financing relative to the equity market.
[11] Statistics on outstanding corporate bonds are obtained from SIFMA 2019 Factbook (http://www.sifma.org/research) and Fleming, Shachar, and Vogt (2015).

bond markets (at least to the extent that the two are integrated).[12] However, as Asness et al. (2013) and Bai et al. (2018) demonstrate, commonly used stock market factors have limited explanatory power with respect to corporate bond returns.[13] The detection of abnormal financial performance is highly model-dependent, and the lack of factor portfolios that effectively estimate bond risk premia hinders both research and investment into alternative areas like ESG (Bai et al., 2018, Bektic, 2018). More generally, when considering any active strategy, investors need to have an understanding of the sources of expected returns, the sustainability of those returns, and the underlying risks. Researchers have recently started to develop factor-mimicking portfolios based on bond characteristics (Bai et al., 2018) and established risk premia (Houweling and van Zundert, 2017, Bektic et al., 2018, Slimane et al., 2018, Israel et al., 2018), but few have incorporated these factors into performance evaluation models.[14]

Although equities and bonds represent claims against the same underlying corporation (Merton, 1974), and investors in each market largely overlap (Sandulescu 2020), we argue that the bond market warrants special consideration due to certain peculiarities of corporate bonds. First, shareholders and bondholders have divergent, and sometimes conflicting, objectives due to their respective positions in a firm's capital structure. While a shareholder's return is unknown and stochastic, that of a bondholder is (mostly) known and constant. So long as a firm can continue to fulfill all of its credit obligations, a worsening business outlook that impacts shareholders does not necessarily effect bondholders (Slimane et al., 2020). As a result, several researchers point out that shareholders should be more sensitive to the extra-financial risks that ESG ratings reflect (Gerard, 2018, Barth et al., 2019). On the other hand, studies increasingly identify downside risk reduction as the principal channel through which ESG influences financial performance (Grossner, 2017, Hoepner, 2018).[15] Since bondholders prioritize the management of credit and default risks, they may be ideally suited to recognize, and discount, the benefits of ESG (Sharfman and Fernando, 2008). Most importantly, studies emphasize ESG's role in either encouraging, or mitigating, agency conflicts between shareholders and bondholders. Hoepner and Nilsson (2020) suggest that excessive investment into ESG may benefit the former at the expense of the latter, and that the risks of asset substitution rise alongside ESG spending. Other researchers argue that ESG signals a lower likelihood that distressed firms will engage in practices that benefit shareholders at the expense of bondholders, such as cash diversion, strategic default, and leveraged acquisition (Bhojraj and Sengupta, 2003, Cremars, 2007, Ferrell et al., 2016, Amiraslani et al., 2019).

Though explored less in this paper, differences between the equity and bond market with respect to the role of credit ratings, institutional investors, investor engagement, and liquidity further highlight the need for a separate investigation of ESG's effect on bond returns. Credit rating agencies increasingly incorporate ESG criteria into their ratings frameworks (Hörter, 2017). Research suggests that credit ratings do not subsume all of the ratings-relevant information contained in ESG scores

[12] There are, however, a number of reasons why equity and bond markets may not be perfectly integrated (see Israel et al., 2018 and Bai et al., 2018). These include bondholders' greater sensitivity to credit and downside risk, lower levels of liquidity in the bond market, the influence of institutional investors, and the (sometimes) conflicting interests of shareholders and bondholders (i.e. leveraged buyouts, the agency costs of debt).
[13] Slimane et al. (2020) finds that, among US corporate bonds, equity factors have greater explanatory power in HY compared to IG (though explanatory power is quite low in both markets).
[14] To our knowledge, the only other paper to do so is Bali et al. (2019), which use the model of Bai et al. (2018) to assess their proposed factor portfolios.
[15] This includes environmental, reputational, litigation, and supply chains risks and scandals (Edmans, 2011, Bauer and Hann, 2014, Seltzer, 2020), as well as ESG's impact on stakeholder trust and cooperation (Lins et al., 2017, Amiraslani et al., 2019), and the insurance-like effects of moral capital (Godfrey et al., 2009).

(Polbennikov et al., 2016, Yang, 2020), but certain commentators expect this to occur in the coming years (Nauman, 2019, Slimane et al., 2020). Second, institutional investors represent a much greater share of all participants in the corporate bond market (86% compared to 64% for equities, Oikonomou et al., 2014). Such investors are in general more sophisticated and long-term oriented, which some argue makes them better suited to recognize the benefits of ESG integration (Waddock and Graves, 1994, Bhojraj and Sengupta, 2003, Dyck et al., 2018). There is also evidence that institutional ownership relates positively to CSP (Brammer and Millington, 2004, and Dyck et al., 2018), and that institutional investors are more likely to avoid socially controversial securities (Chava, 2014, Fernando et al., 2017, Hong and 2020) and to consider climate risks important investment risks (Krueger et al., 2019).[16] Oikonomou et al. (2014) also note that the percentage of institutional investors is negatively associated with the number of free float bonds, which translates into a greater capacity for bondholders to "discipline" firm management. As firms tap the bond market more frequently than they do the equity market (primarily due to the maturing nature of bonds), they are, in general, more exposed to debt market discipline (Salvi et al., 2019).[17] Lastly, the corporate bond market is substantially less liquid than the equity market (Bektic, 2018), which may hinder the efficiency with which the market discounts ESG-related information, and may also facilitate the segmentation required for certain demand-based idiosyncrasies to manifest in bond prices (i.e. the shunned stock effect, Derwall, 2011).

This thesis contributes to the growing strands of literature on both factor and ESG-based investing in the corporate bond market. We develop a cross-sectionally comparable sample of liquid corporate bonds from 1950 U.S. based issuers over the period from January, 2015 to March, 2021. From this we construct mutually exclusive portfolios based on a variety of factor definitions and ESG ratings, and it is along these two categories that we divide our empirical analysis. In the first part, we use the sorted portfolio approach of Fama and French (1992) to examine bond-specific variations of traditional factors such as Value, Size, and Momentum (Asness et al., 2014, Houweling and van Zundert, 2017, Israel et al., 2018, Slimane et al., 2019), as well as those of more recently developed factors like Low-risk, Downside-risk, and Credit Quality (Bai et al., 2018, Bali et al., 2019). In line with evidence that investment grade (IG) and high yield (HY) bonds comprise distinct markets (Ambastha et al., 2010, Chen et al., 2014), we divide our sample accordingly and create two sets of characteristic-sorted portfolios.[18] Like Houweling and van Zundert (2017), we show that a number of factor-mimicking portfolios produce strong returns, which conventional multifactor models fail to explain (such as those from Fama and French, 1993, Elton et al., 1993). We also note some important distinctions between the two markets. Size and Low-risk only produce statistically significant alphas in IG and HY respectively, and while we find evidence of a strong momentum effect in HY, we detect a pronounced reversal effect in IG.[19] Our primary objective, however, is to combine these factors into alternative models with which to analyze the performance of ESG-sorted portfolios. To do so, we examine the relationships between factors and employ an orthogonalization procedure to extract unwanted market and factor exposure (Elton et al., 1993, Hoepner and Nilsson, 2018). Ultimately, we create three models, one of which exclusively includes conventional stock and bond market factors, while the other

[16] More recently, Gillian et al. (2020) call the influence of institutional investors into question, and fail to find any relationship between institutional trading and CSP.

[17] This channel, however, may be double-edged sword, as it potentially promotes the transmission of both value-relevant information and investors' non-pecuniary preferences.

[18] Houweling and van Zundert (2017) and Bektic (2018) do the same, but the majority do not split their samples between IG and HY bonds to conduct their primary analysis (e.g. Bai et al., 2018, Israel et al., 2018, Bali et al., 2019).

[19] This is consistent with the findings of Pepsi and Zhang (2010) and Houweling and van Zundert (2017).

two are comprised of our bond-specific factor portfolios.[20] We find that the latter models capture the variation in the returns to ESG-integrated portfolios more effectively. After Bali et al. (2019), we are the second study to use alternative multifactor models to measure the performance of characteristic-sorted bond portfolios, and we are the first to do so in regards to ESG.

Similar to the literature on ESG and equities (Breedt et al., 2018, Dorfleitner et al., 2020), we present evidence of an ambiguous effect of ESG integration on the returns of corporate bond portfolios. One of our unique contributions to the generally sparse literature concerns the distinct behaviors of IG and HY portfolios. We find that IG portfolios with the worst average ESG ratings generate positive abnormal returns, which remain impervious to controls for prominent bond characteristics and risk factors. Moreover, we use both expanding and rolling windows to show that the performances of the worst-rated portfolios steadily improve over our sample period. In HY, both top- and bottom-rated portfolios tend to underperform mid-rated portfolios, and those comprised of bonds with ratings in the top 20%-40% generate robust and statistically significant alphas. The performance of ESG quintiles in HY, however, appear more sensitive to changes in market conditions, and the alpha of the top-rated Governance portfolio even changes from negative to positive during the Covid-induced crash in 2020. Although our primary ESG ratings come from Thomson Reuters/Refinitiv, we are able to replicate most of our results with the Bloomberg ESG score. We also examine the performance of portfolios based on alternative ESG indicators, including the year-over-year percentage change in ESG ratings (ESG momentum) and the intensity of ESG controversies (Thomson Reuters' news-based controversies rating). In contrast to our IG quintile portfolios based on the *level* of ESG ratings, we find that both top and bottom quintiles with respect to ESG momentum outperform their peers. In HY, we discover noteworthy results after we control for the potential influence of bond size on the effect of ESG controversies. While the latter have a minimal impact on the returns to portfolios comprised of the 50% largest bonds, they significantly impair those of portfolios that contain the smallest bonds. Lastly, we test exclusionary and best-in-class strategies, and, like studies concerning equities (Derwall and Koedjik, 2009, Statman and Glushkov, 2009, Halbritter and Dorfleitner, 2015), we find that ESG investment strategies can be implemented in corporate bond portfolios without an adverse effect on financial performance.

In line with other empirical studies that analyze portfolios constructed around ESG criteria (Hoepner and Nilsson, 2020, Dorfleitner et al., 2020), we propose that the patterns we observe in the returns to ESG-sorted portfolios provide insights into the fundamental relationships between ESG and bond value. Our results in IG are consistent with the predictions of the errors-in-expectations and shunned-stock hypotheses (Derwall et al., 2011).[21] The former suggests that investors struggle to discount the complex financial implications of ESG, and that abnormal returns associated with ESG fade over time as investors learn to value its effects properly (Pereira, 2019). If we assume that bonds from issuers with high ESG ratings outperformed those with poor ratings in earlier periods as the literature suggests (Derwall and Koedjik, 2009, Henke, 2016, Polbennikov et al., 2016), the unremarkable returns of our top ESG portfolios indicates that bond prices now reflect the benefits to strong ESG. The shunned-stock effect asserts that investors' non-pecuniary preferences drive up the expected returns of low ESG securities (Derwall et al., 2011).[22] Since only bonds from the worst-rated issuers display

[20] In both sets of models, we replace the traditional DEF factor (for which most studies use an aggregate bond index) with an equally weighted portfolio for all IG or HY bonds in our sample.

[21] We are the second study to apply these the theories to corporate bond returns, and the first to do so in regards to U.S. bonds (Pereira, 2019).

[22] This so-called shunned-stock effect is based on the Merton (1987) incomplete information model, such that it is possible for investor preferences to limit risk-sharing and therefore to segment markets (in this case between high and low ESG-rated securities).

abnormal returns across each of our ESG portfolios, and since these returns increase over our sample period, we argue that our results fit a demand-based explanation better than they do one related to an ESG risk premium (Pastor and Stambaugh, 2020). We attribute the differences we observe between the returns of our IG and HY portfolios to market segmentation (Ambastha et al., 2010, Jostova et al., 2013, Chen et al., 2014) and the variable relevance of credit risks, especially those concerning agency frictions between shareholders and bondholders (Cremars, 2007, Amiraslani et al., 2019). To explain the underperformance of top-rated HY portfolios, we refer to Hoepner and Nilsson (2020), who argue that risks of asset substitution rise alongside ESG spending, and Dorfleitner et al. (2020), who suggest that the market penalizes small firms for high levels of ESG spending more so than large firms.[23] On the other hand, strong ESG practices help firms to cultivate trust with their stakeholders over time (Porter and Kramer, 1997), which reduces agency conflicts and promotes stability when overall trust is low (Lins et al., 2017, Amiraslani et al., 2019). The literature also demonstrates how ESG management reduces reputational, regulatory, environmental, and litigation risks, which, especially in vulnerable firms, can threaten a firm's solvency (Bauer and Hann, 2014). For these reasons, portfolios of HY bonds from issuers with ratings in the second quintile perform best, while those with the lowest ratings fail to match the returns of their IG counterparts.

In the following section we review the literature with respect to factor-based investing in fixed income, while section 3 includes a review of literature on ESG. Section 4 describes our data collection and filtering process and Section 5 outlines our methodology with respect to the construction and analysis of factor and ESG portfolios. Section 6 includes our results and analysis of factor and ESG-based strategies and Section 7 concludes.

[23] They attribute this to small firms' lower level of scarce internal resources. Not only does this line of reasoning largely apply to HY firms as well, but due to higher required rates of returns, spending on ESG is more costly for HY issuers in general.

2. Literature Review Factor Investing and Corporate Bond Returns

2.1 Brief History of Factor Investing

The origins of factor investing are largely academic and span decades. In his seminal 1952 paper, Markowitz introduced Modern Portfolio Theory, which stresses the principle of diversification to maximize expected returns for a given level of risk. Markowitz conceptualized risk in terms of asset price variance, such that, in an efficient market in which trade prices provide investors with all necessary information, price-return variance accurately reflects risk. At roughly the same time, a group of researchers (Treynor (1961), Sharpe (1965), Lintner (1965) and Mossin (1966)) attempted to explain this variance with the capital-asset pricing model (CAPM). CAPM decomposes risk into a common, market component and a residual, asset-specific factor. Since investors can hedge against idiosyncratic risk through diversification, investors only require compensation for systematic risk. CAPM therefore suggests that the market is the sole relevant factor, and that an asset's co-movement with the market portfolio (an asset's beta) determines its risk premium. As a result, the expected return on any stock can be described as a function of its covariance with the market portfolio.[24]

$$[R_i] - R_f = \beta_i^{market}(\mathrm{E}[R_{market}] - R_f)$$

Where R_i, R_f and R_{market} are the return to $asset_i$, the risk-free rate and the market return respectively, and β_i^{market} is calculated as

$$\beta_i^{market} = \frac{cov(R_i, R_{market})}{\sigma^2(R_{market})}$$

Additionally, the model assumes that investors have homogenous beliefs and that the set of means, volatilities and correlations are the same for all investors. Since CAPM asserts that the market portfolio is the optimal portfolio investors can hold in equilibrium, and because it assumes that everyone has the same mean-variance utility, the market factor becomes the mean-variance efficient portfolio in equilibrium.

The CAPM helped to usher in a new era for portfolio management, providing investors with the first theoretically sound benchmark for returns and reconceptualizing risk. Regarding the latter, researchers increasingly recognized that it was not just an asset's own volatility that determined its risk, but rather its behavior relative to other assets and to the broader market (Ang, 2009). The model advanced the notion that a risk premium is systematic and pervasive, such that it impacts all assets and cannot be arbitraged away. In the case of the market portfolio, this is because the risk premium is a function of underlying investors' preferences. Another way to conceptualize risk according to CAPM is as compensation for losses during bad states of the market. Stocks with higher market betas lose more value during bad states (and gain more value during positive states) compared to stocks with lower betas. Assuming that investors are generally risk averse, gains during positive states of the market do not offset losses during bad states, which means that high beta assets require higher expected returns to be held in equilibrium.

[24] Which can be constructed by holding all securities in proportion to their market capitalization (e.g. a market index fund) such that nonfactor (idiosyncratic) risk is diversified away. This process, according to CAPM, increases risk-adjusted expected returns, as investors are not compensated for nonfactor risk.

Despite the significance of the intuition underlying CAPM, the model's predictions have been famously inaccurate (see Ang, 2014, for a detailed review).[25] Not long after CAPM's introduction, researchers expressed skepticism that a single factor could adequately reflect the complexity of financial markets (Sharpe, 1970, Rosenberg, 1974). They noted that stock returns not only deviated from CAPM-predicted returns, but that they did so in seemingly predictable ways, suggesting the existence of additional, common factors in stock returns (Siegel, 2003). That a single risk factor does not entirely capture the systematic risks investors face underlies Ross's (1976) arbitrage pricing theory (APT). APT borrows heavily from the framework CAPM provides. It is a one-period model based on the absence of arbitrage, such that asset returns are a function of non-diversifiable risks consistent with a factor structure.

$$R_i = \alpha_i + \sum_{j=1}^{n_\theta} \beta_i^j \theta_j + \epsilon_i$$

Where β is the sensitivity of asset i to factor j, θ is factor j, and ϵ is the residual, asset-specific risk of asset i, such that $\mathrm{E}[\epsilon] = 0$, $cov(\epsilon_i, \epsilon_k) = 0$ for i not equal to k, and $cov(\epsilon, \theta) = 0$. More specifically, given that equilibrium prices offer no opportunities for arbitrage, APT, like CAPM, asserts a linear relationship between assets' expected returns and their covariance with systematic factors.

Unlike CAPM, however, APT does not say anything more about the nature of these factors. Instead, APT characterizes factors from a statistical perspective, and simply defines them as variables that explain asset returns. It uses a pricing kernel (or stochastic discount factor) to capture a composite of systematic factors for which investors require compensation. In this way, APT legitimized the search for factors in the cross-section of returns, an exercise that would become an integral part of empirical asset pricing (for detailed information on the APT see Campbell, Lo and MaKinlay (1997) or Huberman and Wang (2005)).

Following APT, researchers have approached factor identification in largely one of two ways.[26] In the first, researchers propose factors based on intuition or economic reasoning and then examine how well these factors explain the cross-sectional variation in estimated expected returns (Huberman and Wang (2005)). For example, referring to the basic dividend discount model, Chen, Roll and Ross (1986) suggest a number of economic variables that should impact expected returns through their influence on discount factors or on the marginal utility of real wealth. They show that changes in the yield curve, expected inflation and industrial production explain a significant portion of the cross-sectional variation in stock returns. In the second approach, researchers begin by searching for persistent and economically interpretable patterns in average returns. For example, beginning in the 1980s, researchers found that stocks from firms with low market capitalization (Banz, 1981), high earnings-to-price ratios (Basu, 1983) or high book-to-market equity ratios (Rosenberg et al., 1985) were associated with high average returns, even after controlling for stocks' exposure to the market factor. In one of the most influential papers on factor investing, Fama and French (1992, 1993) examine the relevance of these so-called "size" and "value" effects in explaining stock returns. The authors group stocks into quintile portfolios based on market capitalization and book-to-market ratios, and show that risk-adjusted returns increase as the former decreases and the latter rises. They then build two factor

[25] See, for instance, Ang (2014) for a review of the many empirical studies that have undermined the validity of CAPM's predictions.

[26] There is arguably a third approach that involves more advanced statistical methods, notably principal components analysis, to find factors that best explain the variation in asset returns (Cochrane and Piazzesi, 2005)

mimicking portfolios, the “small-minus-big” (SMB) and “high-minus-low” (HML) portfolios, which fund long positions in small capitalization stocks (high book-to-market) with short positions in large capitalization stocks (low book-to-market) after sorting on book-to-market (market capitalization). Using time series and cross-sectional regression analysis, they show that, while the market portfolio alone explains approximately 70% of the variation in the returns to diversified portfolios, the inclusion of SMB and HML portfolios result in a model capable of capturing up to 95% of the variation in U.S. stock returns from 1966 to 1990.

This result had two profound effects on empirical asset pricing. In the equities literature, Fama and French (1993) motivated thousands of researchers to construct their own factor mimicking portfolios in an attempt to capture as of yet unexplained variation in stock returns. Based on evidence that recent stock winners tend to out-perform recent losers, Jegadeesh and Titman (1993) construct the momentum (UMD) factor. They find that a long-short portfolio of U.S. stocks that buys winners from the past 3-12 months and shorts losers generates abnormal returns that the MKT, SMB and HML factors fail to explain. This result was so striking that Carhart (1997) proposed a four-factor model, which adds momentum as an additional explanatory variable to the Fama-French three factor model. Similarly, Frazzini and Pederson (2014) develop the betting-against-beta (BAB) factor based on the observation that, in direct opposition to CAPM’s prediction that riskier stocks earn higher returns, the least risky stocks tend to generate above-market returns (Blitz and Vliet, 2007). The authors show that a beta-neutral, long-short portfolio that goes long low-beta stocks and short high-beta stocks generates abnormal returns and exhibits a low correlation to well-established stock market factors. Lastly, Fama and French have revisited their model many times (Fama and French 1995, Fama and French, 2008), eventually expanding it to include factors capturing investment growth and profitability (Fama and French (2015)).

Fama and French (1993) also helped to reshape portfolio evaluation and construction. A simple time-series regression of a hedge fund manager’s returns against those of the market and factor-mimicking portfolios that comprise the Fama and French three-factor model reveals the extent to which that manager’s returns can be ascribed to exposure to systematic risks or to a residual (potentially skill-related) component embedded in the regression’s intercept (or alpha). The implication of zero-alphas with respect to factor portfolios provides a strong theoretical justification for factor-based investing. As Martenelli (2015) notes, it can be shown that a theoretical single-step solution to a mean-variance optimization problem coincides with an optimal linear combination of mean-variance efficient benchmark portfolios investing in securities that each have a zero alpha when regressed against the benchmark portfolios. The implication, therefore, is that the most effective way to group securities is not to form arbitrary asset class indexes, but instead to form factor indexes, or factor mimicking portfolios that can be collectively considered as linear proxies for the unobservable stochastic discount factor (Martenelli (2015)). Ang, Goetzmann and Schaefer (2009) made this exact argument in their highly influential analysis of the performance of the Norwegian Government Pension Fund (NBIM) before and during the 2008 financial crisis. The authors showed that well-established pricing factors (including *MKT, SMB, HML* and *UMD*) explained the majority of NBIM’s active returns. This study marked a turning point in factor-based investing, especially since traditional asset allocation strategies and active management had failed to provide adequate downside protection during the global financial crisis (Blyth, Szigety, and Xia, 2016). What had primarily existed in the academic realm became an integral part of investment management, providing practitioners with a more effective way to evaluate and allocate portfolios.

2.2 Fixed income asset pricing

In contrast to the quantity of research on factor-based strategies in equity markets, factor-based investing in fixed income has been slow to develop and remains in an immature state. This is consistent with the general lack of research on bond pricing, especially in regards to corporate bonds.[27] For example, while Chen et al. (1986) and Fama and French (1992) use the dividend discount model as a foundation to explain why the variables they propose relate to average returns, there is no agreed upon framework to estimate credit returns (Israel et al. 2018). As a result, there is no consensus among researchers or investors on how to measure the financial performance of corporate bond portfolios, which we argue remains one of the biggest hurdles facing any researcher seeking to analyse synthetic bond portfolios.

This can be at least partially attributed to three related factors. First, there is a comparative lack of data and transparency regarding bond pricing (especially prior to the introduction of TRACE in 2002). Second, the construction of a single index model akin to the CAPM in fixed income markets poses additional challenges. The CAPM was crucial to development of asset pricing in equity markets, and although Sharpe (1964) does not state explicitly that the assets his model refers to are equity shares, the majority of early research building on the CAPM focuses on NYSE stocks (Black et al., 1972, Fama and MacBeth, 1976). Roll (1977) argues that a cap-weighted combination of all available bonds should represent the efficient set within the fixed income class. To construct an all-inclusive and stable benchmark in fixed income, however, is far more challenging than it is in the equity market. This is due in part to the fact that bonds mature, and that the bond market is primarily a dealer's market, making many bonds, especially corporate issues, very illiquid relative to equities. As a result, bonds are notoriously difficult to short (Asquith et al., 2013) and trading costs of bonds are very high relative to their underlying volatility (Bessembinder et al., 2006). Siegel (2003) questions whether Roll's proposition is even theoretically justifiable given the presence of offsetting claims in the corporate bond market (particularly in structured debt and derivatives) and the lack of clarity about whether government bonds constitute wealth in the real economy. Furthermore, while researchers have criticized value-weighted portfolios in equity markets (Fama and French (2008)), Siegel argues that this practice is far more dubious in corporate bond markets as its impact on the distribution of risk is more severe (the so-called "bums problem").[28] Aside from the challenges in creating a stable benchmark against which to assess bond returns, however, is the question of whether a single index model is even appropriate. This relates to the third major impediment, which concerns fragmentation across fixed income markets. Researchers show that bond characteristics such as option features (Elton et al. 2001) or credit ratings (Chen et al. 2011) segment fixed income markets. Hoepner and Nilsson (2018) note that this segmentation precludes researchers from relying on single index models to price most bond portfolios without risking significant misspecification biases. What little research has emerged regarding the financial performance of fixed income funds and securities, therefore, relies primarily on multi-index models to capture the differences in exposure of various bond classes (Elton et al. 1995, 2001, Derwall and Koedijk, 2009, Hoepner and Nilsson, 2018).

To assess the performance of bond funds and investment strategies, researchers depend on variations of one of two basic models. The first, that of Fama and French (1993), combines three equity market factors (the market portfolio and the two factor mimicking portfolios, SMB and HML) with two

[27] Sovereign bond risk modelling received serious attention much earlier than did corporate bonds. See Fama and Bliss (1987), Nelson and Siegel (1987), Litterman and Scheinkman (1991) and Cochrane and Piazzesi (2005)).
[28] Companies with larger market capitalizations are generally also more diversified and face lower default risk, thus causing large bond issues to have a higher probability of being of a lower credit quality.

bond market factors constructed from fixed income indices. The authors reference Merton (1974), as do many subsequent researchers, to justify the application of equity market factors to explain bond market returns. The second model, that of Elton et al. (1995), is a multi-index model that drops SMB and HML and adds factors to capture the exposure of a security or a portfolio to other bond classes. As this section details, there are clear disadvantages to both approaches, but in the absence of an agreed upon framework with which to measure expected corporate bond returns, they remain the basis for most fixed income pricing models in the literature. In the past few years, however, a handful of researchers have started to examine characteristic-sorted portfolios in fixed income (Houweling and van Zundert, 2016, Bai et al., 2018, Israel et al., 2018, Bali et al., 2019). In the last part of this section, we, thus, describe how we borrow from these studies to develop more robust models with which to measure the performance of ESG-sorted portfolios and strategies.

2.2.1 Equity market factors

Merton's (1974) structural model of credit risk, as well as its many variants (cite), have become some of the primary workhorse models in pricing corporate bonds. Merton asserts that since both holders of equity and of debt own claims against the same underlying corporation, fundamentals that drive stock prices should also impact credit spreads in order to ensure the absence of arbitrage. He employs a standard option-pricing framework with which it is possible to derive the value of credit as simply the value of a European put option on the firm's assets with an exercise price equal to the face value of the firm's total outstanding debt. Similarly, treating all of a firm's debt as a single, zero-coupon bond, the value of a firm's equity becomes equal to the value of a call option on the underlying firm assets with a strike price equal to the firm's debt. In this case, at the bond's maturity date, shareholders will abandon their investment should the firm's value fall below the value of its debt. If a firm's value is greater than the face value of its debt, however, shareholders will exercise the option and collect the difference. The difference between firm value and the value of outstanding debt is referred to as the "distance to default". Assuming the frequency of default follows a normal distribution, ranking firms along this measure of creditworthiness provides an intuitive estimate for the probability of default.

Merton's model has had two very important and related implications that have shaped studies on empirical bond pricing. First, it predicts that an issuers' credit risk is the main factor driving yield spreads, and second, it links a firm's credit risk to its capital structure, such that the price of a bond is tied directly to the price of a corresponding equity share. The Merton model and its extensions, however, have at best a mixed track record in pricing corporate bonds. Only a few years after Merton published his paper, Jones, Mason and Rosenfeld (1984) demonstrated that his model significantly under-predicted observed corporate bond spreads, a result that many papers have since replicated (Leland and Toft, 1996, Anderson and Sundaresan, 1996). Other researchers, such as Collin-Dufrense and Goldstein (2001) and Huang and Huang (2003), find no fault with the efficacy of structural models in pricing default, and instead suggest that corporate bonds are influenced by factors unrelated to credit risk and therefore absent from structural models altogether. Researchers have proposed additional variables, including ESG factors, to help explain this so-called "credit spread puzzle", but there is no consensus as to what model or combination of factors could resolve it (see Elton et al. 2001, Longstaff et al. 2005, Menz, 2010, Chava, 2014, Feldhütter and Schaefer, 2018 and Bai et al., 2020). Despite the inability of structural models to explain changes in credit spreads, many researcher reference Merton (1974) to explain the integration between debt and equity markets, and to justify the application of equity pricing models to the corporate bond market.[29]

In their seminal 1993 paper, Fama and French suggest that a single model should be sufficient to explain the returns to both stocks and bonds should the two markets be integrated. The authors find, however, the explanatory power of their three-factor model to be lacking in regards to corporate bond returns. They therefore supplement their three-factor model with two bond specific factors, *TERM* and *DEF*, designed to capture maturity risk and default risk respectively. The authors argue that since the risks associated with unexpected changes in interest rates and in the probability of default are common to all bonds, they should be rewarded. Given their higher exposure to interest-rate risk, bonds with longer durations should, on average, earn higher returns than short-term bonds. Similarly, bonds that bear default risk should command a higher premium than those that do not, especially since default events are more likely to occur during economic downturns.

They measure the *TERM* factor as the returns of a long-term government bond index in excess of the one-month Treasury bill and the *DEF* factor as the excess returns of a long-term corporate bond index over those of a long-term government bond index. The authors find that, within their sample, *TERM* and *DEF* appear to be common factors in bond returns. More specifically, they find that all of their bond portfolios, irrespective of credit rating, exhibit statistically significant loadings on the two factors. Additionally, they show that what little variation the equity market factors explain in both the time series and the cross-section of corporate bond returns vanishes almost entirely when they introduce *TERM* and *DEF,* such that the bond market factors render the three-factor model redundant. On the other hand, they show that, even after controlling for stock market factors, the bond market factors explain a non-trivial portion of the time series variation in stock market returns.[30] Ultimately, while they find that their five-factor model explains as much as 95% of the in-sample return variation of U.S. stocks, the same model explains only 79% of the variation in high-grade corporate bond portfolios, and less than 49% of the variation in low-grade U.S. corporate bonds.

Despite the comparatively mixed results of the Fama and French five-factor model in explaining corporate bond returns, it has persisted as one of the core models to evaluate the performance of bond portfolios. Many of the most influential papers in regards to the recent push to develop bond-specific factor models rely on equity market factors to assess their factor-mimicking portfolios (Houweling and van Zundert 2017, Bektic et al. 2017, Bai et al., 2018, Israel et al., 2018, Bali et al., 2019). Bai et al. (2018) note the apparent advantages of this approach, such that researchers have a theoretical justification for borrowing from a much more developed body of literature. Insofar as discovered equity market factors represent rational pricing factors, risk premiums estimated in the equity market should be consistent with those in the corporate bond market. The extent to which the two markets are integrated, however, is still hotly debated, with some referencing the failure of structural models of credit risk in order to emphasize market segmentation and the resulting need to develop bond-specific factors (Kapadia and Pu (2012), Israel et al. (2018). On the other hand, Schaefer and Strebuleav (2008) find that credit risk models accurately predict empirical hedge ratios between stocks

[29] It should be noted that, due primarily to the greater popularity of research into equities, researchers primarily invoke Merton to use bond market variables to explain equity returns. Vassalou and Xing (2004) show that their measure of default risk (derived from Merton (1974)) explains cross-sectional equity returns and helps to account for the size and value effects. Campello et al. (2008) similarly refer to Merton (1974) to justify their use of bond yields to proxy for ex-ante equity returns in order to assess the Carhart (1997) four-factor model.

[30] In contrast to Chen, Roll and Ross (1986), however, Fama and French note that the average premiums on TERM and DEF are too small to explain much variation in the cross-section of average stock returns.

and bonds, and Bektic (2018) notes that the growth of CDS markets and the popularity of capital market arbitrage (Duarte et al. (2007) have likely increased the link between equity and bond prices. What seems incontrovertible, however, is the poor ability of factor-mimicking portfolios constructed from equity securities to explain bond returns, with many researchers finding that such portfolios explain little to none of the time series variation in corporate bond returns (Asness et al., 2015, Bai et al., 2018). Additionally, Asness et al., (2015) provide evidence that characteristic portfolios from one asset class exhibit low to no correlations with those from other asset classes. In line with these results, we show that value, size and momentum portfolios constructed from individual bonds have limited (and sometimes even negative) correlations with corresponding equity market portfolios.[31]

2.2.2 Multi-index models

In response to what was already a very clear underrepresentation of fixed income in factor-based research, Elton et al. (1995) sought to create a fixed income asset-pricing model that could explain the returns of a range of different bond funds. They developed their model with a sample of 179 U.S. bond indices, bond funds and bond mutual funds over the period 1986 to 1991. Their resultant four-factor model has become the most cited asset-pricing model for fixed income and time series analysis (Hoepner and Nilsson, 2018).

$$R_{p,t} - R_{f,t} = \alpha_i + \beta_{1,i}(Market_{m,t} - R_{f,t}) + \beta_{2,i}Default_t + \beta_{3,i}Option_t + \beta_{4,i}Equity + \varepsilon_{i,t}$$

Their Market factor, constructed as the excess return of a broad bond market index, is intended to capture a security or portfolio's exposure to the investment grade bond market. Modeled after Fama and French's *DEF* factor, the Default factor is constructed as the excess return of a high-yield corporate bond index and, in covering exposure to the high yield market, is supposed to represent default risk. The Option factor, which measures the return of a mortgage-backed security index over an intermediate-term government bond index, is included to capture exposure to securitized debt and to account for a bond's option characteristics.[32] Lastly, the Equity factor, simply constructed as the excess return of an equity index, covers exposure to the equity market. Elton et al. (1995) also considered the *TERM* factor of Fama and French (1993) as well as two macroeconomic factors measuring unanticipated changes in inflation and changes in GNP, but found that they did not contribute significantly to their model.[33] A wide array of papers, primarily those investigating the returns of bond funds and mutual funds, have either utilized the four-factor model or constructed similar multi-index models based on the findings of Elton et al. (1995).[34] Hoepner and Nilsson (2018) investigate the four-factor model, as well as the many modifications researchers have made to it since its inception, against a large sample of 650 US and European fixed income funds from 2000 to 2016. Based on their results, the authors propose an enhanced, ten-factor model, which captures well over 90% of the time-series variation of the bonds funds in their sample.

---

[31] Israel et al. (2018) and Bektic (2018) show a similar phenomenon with their respective factor mimicking bond portfolios.

[32] Given uncertainty inherent in the freedom homeowners have to either refinance or sell their homes at almost any time, mortgage-backed securities have distinct, option-like characteristics.

[33] Hoepner and Nilsson (2018) note that subsequent papers have found these two macroeconomic variables to be more important in cross-sectional analysis than in time-series analysis.

[34] Some include Huij and Derwall (2008), Gutierrez et al. (2009), Lipton and Kish (2010), Cici and Gibson (2012) and Henke (2016). See Hoepner and Nilsson (2018) for a more detailed review.

$$R_{p,t} - R_{f,t} = \alpha_i + \beta_{1,i}(Market_{m,t} - R_{f,t}) + \beta_{2,i}Default_t + \beta_{3,i}Option_t + \beta_{4,i}Equity_t + \beta_{5,i}Duration_t + \beta_{6,i}Global_t + \beta_{7,i}Credit_t + \beta_{8,i}€_t + \beta_{9,i}£_t + \beta_{10,i}¥_t + \varepsilon_{i,t}$$

The Market, Default, Option and Equity factors represent the factors in the base four-factor model. They construct the Duration factor, which they intend to cover interest-rate risk like the *TERM* factor, as the return difference between a long-term and short-term government bond index.[35] To take into account a bond fund's exposure to the global market, the authors also add a Global factor measured as the excess return of a broad global investment grade index. The authors follow the orthogonalization approach used in Elton et al. (1993) to correct for the high correlation of the Global factor with the Market factor (we describe this procedure in section 4). The Credit factor is the same as the *DEF* factor, which the authors measure as the return of a long-term corporate bond index over a long-term treasury index. They argue that the Credit factor more accurately captures exposure to default risk than the Default factor alone. Lastly, the authors include three exchange rate factors, consisting of the three most traded currencies after the USD, all quoted in USD (EUR:USD, GBP:USD and YEN:USD).

Given fragmentation across fixed income markets, multi-index models like the Elton four-factor model and the expanded ten-factor model of Hoepner and Nilsson (2018) are among the best tools with which to evaluate the performance of corporate bond funds and mutual funds, especially where the nature of the holdings are not precisely known. In their investigation of SRI bond funds, Derwall and Koedjick (2009), Henke (2016) and Leite and Cortez (2016) use variations of the Elton four-factor model to determine whether sustainability factors translate into an outperformance of SRI funds relative to conventional funds.[36] However, lacking many alternatives, researchers continue to rely on multi-index models in combination with equity factor portfolios to measure the performance of synthetic bond portfolios as well. For example, the only two studies that investigate the role of ESG-factors in corporate bond portfolios (Perreira et al., 2019, and Hoepner and Nilsson, 2020), both use derivations of the Elton model.

2.2.3 Alternative corporate bond factors

Bai et al. (2018) highlight the danger of using an model that fails to adequately account for exposure to systematic risks. The authors show that, when using multi-factor models constructed from equity market factors and bond indices, the majority of their test portfolios exhibit seemingly strong, statistically significant alphas. Upon replacing these models with a factor model based on characteristic-sorted bond portfolios, however, they find that many of these alphas disappear entirely.

Over the past few years, researchers have started to apply methods predominantly used in the equities literature to develop factor models in fixed income. The first of these studies, which Slimane et al. (2019) identify as the "veritable turning point" in the literature on factor investing in fixed income, is that of Houweling and van Zundert (2014, 2017). The authors use decile analysis to construct factor-mimicking portfolios based on well-established factors such as Size, Value, Momentum and Low-risk. The authors show that these portfolios generate significant risk-adjusted returns in-sample, that these returns are robust to a range of alternative specifications including different factor definitions, and that the various factor portfolios exhibit low correlations with each other. Slimane et al. (2019) replicate these results with slightly modified factor definitions, and additionally

[35] This is essentially the same as what some authors refer to as the *TSY* factor (Israel et al., 2018.

[36] Leite and Cortez also integrate conditioning information to allow for time-varying alphas and betas.

show that these factor portfolios explain a significant portion of the cross-sectional variation in corporate bond returns. Israel et al. (2018) propose Value, Defensive, Momentum and Carry portfolios using definitions that are similar to those of Houweling and van Zundert (2017). They show that these factors, together with controls for variables such as rating and maturity, explain up to 15% of the cross-sectional variation in bond returns, a phenomenal result in comparison to even the most recent research into the cross-sectional determinants of equity returns.[37] Bektic (2018) applies the equity market definitions of Size, Value and Momentum (defined in terms of equity market capitalization, book-to-market and equity price momentum respectively) to construct factor mimicking portfolios with corporate bonds that generate strong within-sample risk-adjusted returns. Bai et al. (2018) closely follows the methodology of Fama and French (1992, 1993) to propose a four-factor model that include proxies for downside risk, credit quality and liquidity risk, as well as a short-run reversal factor based on past one-month returns. Bali et al. (2019) examine a momentum factor, along with a long and short-term reversal factor, which they show contributes significantly to models that include both traditional equity and bond factors, as well as to the four-factor model of Bai et al. (2018). We borrow from these papers to construct alternative, factor-mimicking portfolios with which to analyze the performance of ESG-based portfolios and strategies.

2.3 Development of an alternative factor model

Among the many considerations we make to determine which factors to include in our model, there are two that require additional clarification. The first concerns data availability and the nature of our bond screening process (which we address in greater detail in section 3). For example, data limitations prevent us from constructing the Liquidity (*LIQ*) factor from Bai et al. (2018), while our bond filter, which selects bonds based on issue date, makes it infeasible for us to emulate the Long-term Reversal (*LTR*) factor of Bali et al. (2019).[38] The second consideration relates to the strength of the economic justification for the existence of each factor. Researchers have started to raise the alarm about the hundreds of supposed factors discovered in the equities literature. Harvey et al. (2016) and Feng et al. (2020) propose more stringent testing to determine whether prospective factors do indeed reflect systematic risk premia, or whether they simply proxy for similar fundamental risk factors (and are therefore redundant) or are statistical aberrations (Chen and Kim, 2019). Data-mining is a legitimate concern, especially given the well-known biases that can emerge from some of the most fundamental factor models, including the Fama-MacBeth (1973) two-pass regression and its variants. In addition to the classic errors-in-variables problem and other well-known issues (see Kim (1995)), research shows that such biases can also arise from something as simple as a researcher's construction of test portfolios (Ang, Liu and Schwarz, 2018). Importantly, that the APT defines factors entirely in statistical terms makes the economic reasoning accompanying factor selection an integral measure to protect against data-mining or the misinterpretation of spurious results (Martenelli, 2015). Replication is also highly important, and although there are researchers who find fault with some of the most well-established factors (for examples of criticisms levied against size and value, see Dichev, 1998, Horowitz et al., 2000, Schwert 2003, Penman et al. 2007 and Lioui and Poncet, 2012) these factors have been

[37] As the authors note, Lewellen (2015) finds that 15 equity characteristics explain 7.6% of the cross-sectional variation of equity return

[38] For reasons we explain in section 4, at one stage of our representative bond filter we retain only the newest bonds available for each issuer (less than two years since issuance date). As a result, we only have 36 months of returns data (which is how Bali et al. (2019) define their long-term reversal factor) for a portion of the bonds in our sample.

tested hundreds of times in a global context and across large swaths of time using high quality data (Ziegler et al. 2007, Artmann et al., 2012, Baker and Haugen, 2012).

Especially given of our relatively narrow sample window of six years, we are less interested in simply choosing among factors that best explain the time-series or cross-sectional variation in-sample, and instead choose from those that are the most well-established in the literature. We avoid factors that researchers show are highly sensitive to their definitions, such as profitability (Crawford et al., 2015) and assume that, although characteristic-sorted portfolios in one asset class have limited explanatory power in another, rational pricing factors are not dependent on any one market structure. Although it is still a novelty in the fixed income literature to borrow and recreate factor-mimicking portfolios, it is not without precedent. Bai et al. (2018) construct a bond Momentum portfolio, which they add to a multi-factor model that includes equity and index-based factors. Bali et al. (2019) borrow the four-factor model of Bai et al. (2018), which they use alongside traditional equity and bond market factors to assess the performance of their momentum and reversal-based portfolios. In the section below, we review the literature behind each of the factors that we select.

2.3.1 Momentum

Over the past decade, researchers have identified a momentum effect in the corporate bond market. The effect, however, appears less pronounced, and is more nuanced, than it is in the equity market. For example, although Gebhart et al. (2005) find that high past equity returns predict high bond returns, they fail to discover a bond-specific momentum effect in investment-grade bonds.[39] Pospisil and Zhang (2010), on the other hand, show that momentum strategies (using 1-24 month look-back periods) generate profits in the high-yield bond market. In the most rigorous analysis to date of momentum in the corporate bond universe, Jostova et al. (2013) affirm the results of the previous papers. Ranking bonds over periods ranging from three to twelve months, the authors find low and statistically insignificant momentum profits among investment grade (IG) bonds, but high and significant returns among non-investment bonds for the period 1973 – 2011. They show that these results are robust to changes in various bond-level characteristics such as age, amount outstanding and duration, and that traditional Fama-French factors (SMB, HML, TERM and DEF) fail to explain the variation in returns to bond momentum strategies. The only variable the authors find to be positively and significantly related to the returns to momentum strategies is the proportion of NIG bonds in a given sample. Subsequent empirical papers corroborate this result (Bali et al. (2019)), with some detecting slight, but statistically significant, momentum effects in the IG market (Houweling et al. (2014), Israel et al. (2018)). Houweling et al. (2017) note that, when including HY bonds in his sample, the momentum factor becomes the most prominent, and the least correlated with traditional pricing factors.

Jostova et al. (2013) investigate the relevance of previous explanations for the momentum effect, including spillovers from equity markets, trading frictions and price underreaction. They find that, in general, issuers with the best performing bonds do not also have the best performing stocks, which they argue makes it unlikely that bond momentum simply reflects equity momentum. Transaction costs and trading frictions are also unlikely to account for bond momentum, since the high credit-risk bonds that tend to drive momentum profits face relatively low effective trading costs and fewer

[39] They even find evidence for a modest, but statistically significant, short-term reversal effect, where the top performing bonds over a three-to-twelve-month period underperform past losers over the subsequent three to twelve months. While later papers confirm that the momentum effect is slight, if it exists at all, in the investment grade universe, few find evidence of a short-term reversal effect (Bali et al., 2019).

frictions due to higher turnover rates and larger trade sizes.[40] Additionally, they argue against market opaqueness as a potential explanation for the persistence of bond momentum, since they find that bond momentum profits increased following the improved market transparency brought on by TRACE in 2002. The authors, however, retain the possibility that the hypothesis of Hong and Stein (1999) regarding the gradual diffusion of information in market prices explains, at least partially, the momentum effect in bonds. This is because they find that momentum profits are highest among both HY bonds and private-firm bonds, where information is harder to interpret and is likely to diffuse more gradually. Lastly, Bali et al. (2019) emphasize the role of default risk in explaining the returns to portfolios sorted on past returns, finding that momentum winners exhibit higher market risk, higher credit risk and higher interest-rate risk. They also find that the momentum effect is stronger during economic contractions and periods of elevated default risks, to the extent that the return spread between momentum winners and momentum losers disappears entirely when they exclude the 2008-2009 period. As a result, they argue that momentum in the bond market is restricted in the cross-section to default-prone bonds and in the time series to crisis periods.

2.3.2 Value

Despite its significance as a pricing factor in equity markets, the value factor has received very little attention in debt markets until recently. This primarily due to the difficulty in finding a substitute for the book-to-market ratio (Israel et al., 2018). The literature identifies two sets of potential substitutes, including past returns (Fama and French (1996) and Asness et al. (2013)) and ratios which compare default probabilities to market price (Correia et al. (2012), Houweling et al. (2014), Israel et al. (2018)). There is a robust strand of literature that examines the former (primarily in equity markets) and finds a strong relationship between the HML factor and a long-term reversal effect (DeBondt and Thaler, 1985, Fama and French, 1996, Gerakos and Linnainmaa, 2012, Asness et al., 2013). Asness et al. (2013) and Bali et al. (2019) detect a strong, long-term reversal effect in bonds, supporting the notion that a similar premium exists in bond markets as well. Correia et al. (2012) attempt to replicate the intrinsic-value-to-price structure of value factors studied in equity markets to create a value signal for the corporate bond market. They compare a several fundamental measures of risk such as leverage, profitability and Merton's distance-to-default to credit spreads, which represent the market's assessment of credit risk. They find that the differences between their forecasts of default probability and the default probability embedded in credit spreads is highly mean reverting, and significantly and negatively associated with a bond's returns over the subsequent six months. Their results are robust across industries and to various research design choices, and traditional equity factors fail to explain the returns to portfolios sorted on bond-specific measures of value. Correia et al. (2012) inspire a number of alternative value proxies in more recent papers. Houweling et al. (2017) use a simple linear regression model to predict credit spreads based on time to maturity and credit rating. They then construct a value signal based on deviations between actual and predicted spreads, such that they go long those with the largest negative residuals. Israel et al. (2018) similarly compare credit spreads to various measures of default risk. They construct two value measures, one based on Merton's distance-to-default, and, to maximize coverage, another measure that combines credit rating, bond duration and the volatility of bond excess returns. The value measures of both Houweling et al. (2014) and Israel et al. (2018) display economically and statistically significant premiums than cannot be explained

[40] Bonds in the worst-rated quintile, Q5, are four times more likely to trade than bonds in the best-rated quintile, Q1. The average trade size of Q5 bonds is twice that of Q1 bonds, which lowers effective transaction costs despite the adverse impact of higher credit risk. Transaction costs represent only 23% of gross momentum profits in NIG bonds and 16% in private-firm NIG bonds.

by traditional equity and bond market factors and is robust across industries, time, and changes in various bond characteristics (unlike the momentum factor, they also find it to be significant in both the investment grade and high yield bond universes).

The two sets of value measures described above both seek to identify cheap bonds, which they define as either a bond trading at a discount to its historical price, or a bond with a wide spread relative to various proxies for default risk. As these are akin to "cheapness" in the equities space, some of the explanations for the value factor in the equity market may also apply. Risk-based explanations for the value factor in equity markets postulate the existence of a systemic risk factor to which "cheap" stocks are more exposed than expensive stocks. For example, Fama and French (1996) suggest that value firms (which tend to be in cyclical industries) have a more difficult time shifting to more profitable activities in recessionary environments compared to growth firms, which are generally less capital intensive and can therefore divest more easily. Due to the pro-cyclical nature of value stocks, they therefore attribute the positive average returns on the value factor to a recession premium. In line with this risk-based explanation, Cooper and Maio (2016) find that the returns on the HML factor forecast changes in the Chicago Fed national activity index. This could be the case in bond markets as well, as Bali et al. (2019) find that default risk largely accounts for the returns to the long-term reversal factor. Lakonishok et al. (1994), however, find that value strategies in equity markets are no risker than their growth-based counterparts, leading them to reject risk-based explanations for the value premium. Instead, they posit that excessive investor optimism about stocks that have performed well in the past, as well as excessive pessimism about those that have performed poorly, better explains the value phenomenon. Given elevated price-pressure effects in the corporate bond market (Dynkin et al. 2002), such behavioural explanations should apply even more so to bonds.

Since we filter bonds based on time since issuance means we cannot capture value with a reversal portfolio, and we therefore borrow from the latter.

2.3.3 Safety/low volatility/low risk

The "low-risk" or "low-volatility" factor traces its origins to one of the first documented anomalies in capital markets (Carvalho et al. (2015)). Early researchers demonstrated that the relationship between risk and return was much flatter than the CAPM predicted. Haugen and Heins (1972) find that in the US equity market between 1926 and 1969, the least risky stocks exhibited significantly greater returns than the CAPM predicted, while the riskiest stocks underperformed the CAPM. Fama and French (1992) confirm this flat empirical relationship between CAPM beta and returns for the period from 1963 to 1990, and researchers have since replicated these results in both the U.S. equity market (Baker, Bradley and Wurgler (2011)) as well as international developed (Frazzini and Pederson (2011)) and emerging equity markets (Blitz, Pang and van Vliet (2013)).

There are several theoretical and behavioral explanations for the low-risk effect, both in general and in fixed income in particular. Black et al. (1972) highlight the impact of leverage constraints, and posit that investors prefer high beta assets to the potentially costly alternative of levering up investments in low-beta assets. As a result, risky high-beta assets require lower risk-adjusted returns than do low-beta assets. The authors also show from a theoretical perspective that, when the cost of leverage exceeds the risk-free rate, the relationship between risk and return must be much flatter than the CAPM predicts. Frazzini and Pederson (2014) also attribute what they describe as the low-beta effect to constraints on leverage and provide empirical evidence that the risk-return relationship flattens as funding constraints tighten. They also examine the holdings of investors who are more (less) leverage constrained to test their prediction that such investors prefer high (low) beta stocks.

Consistent with their prediction, they find that the beta of the average stock held by leverage-constrained equity mutual funds is 1.08.[41]

While leverage constraints help explain why low-risk stocks appear underpriced, short-selling constraints potentially provide another explanation for why high-risk stocks may be overpriced (Blitz et al. (2013)). Blitz et al. cite the winner's curse phenomenon outlined in Miller (1977), which suggests that in a world with little to no short selling, optimistic expectations will artificially drive up the prices of specific securities. They argue that high-risk stocks are more likely to be overpriced than low-risk stocks, a view supported by Hong and Sraer (2012) who show that high-beta assets are more sensitive to a divergence-of-opinion about their payoffs, making them more prone to speculative over-pricing compared to low-beta stocks. In other words, short-selling constraints prevent investors from correcting the inflated prices of high-volatility securities.

As Houwleing and van Zundert (2014) point out, explanations related to human behaviour, incentive structures and constraints should apply to the corporate bond market as well, perhaps even more so given that investors in the bond market face even greater leverage and short-selling constraints. Carvalho et al. (2015) note the presence of a low-risk/low-volatility anomaly across a broad range of fixed income assets, including high-yield and investment grade corporate bonds. Concerning the low-risk anomaly in the USD corporate bond market, the results of Ng and Phelps (2015) are less supportive and suggest that the apparent anomaly is sensitive to the selected measure of risk. Most notably, asset price volatility and beta, which is how research on equities define low risk, are not necessarily appropriate measures in the corporate bond market (as the volatility of a bond shrinks to zero as it approaches maturity (Israel et al. (2018)). Researchers, however, have suggested a range of alternative measures, the two most promising of which include maturity and credit quality. Ilmanen et al. (2004), Derwall et al. (2009) and Binsbergen and Koijen (2015) document that short-duration bonds outperform long-duration bonds on a risk-adjusted basis, while Kozhemiakin (2007) and Frazzini and Pedersen (2014) find that highly rated bonds outperform low-rated bonds. Ilmanen (2011) and Houweling and Zundert (2014) create a low-risk factor that combines duration and maturity and show that this factor generates strong risk-adjusted returns, which Slimane et al. (2019) find explains a significant portion of time series variation in bond returns beyond that which is captured by traditional factors. Israel et al. (2018) arrive at similar results with their "defensive" factor, which they create by combining duration with gross profitability and market leverage.

2.3.4 Size, Downside Risk and Credit Quality

There has been a lot of research on a Size effect in equity markets related to a company's market capitalization (van Dijk, 2011). Small companies are generally associated with lower liquidity and higher default risk, and should therefore outperform large firms to compensate investors for the additional risks (Banz, 1981). Unfortunately, evidence of a Size effect in debt markets, especially with bond-specific measures such as total debt capital outstanding or bond size, are limited to a handful of studies (Houweling and van Zundert (2017), Slimane et al. (2019)). Houweling and van Zundert (2017) show that find evidence of a size effect with respect to an issuer's total amount of debt outstanding, but offer no economic explanation as to why the effect should persist. Due to data constraints we have to rely on an even less tested measure of size, issue size, which Houweling and van Zundert (2017) show performs similarly to their primary measure. We also find that, across our entire sample, 32% of investment grade issuers and nearly 70% of high yield issuers have only one bond, which

[41] See Frazzini and Pederson (2014) table 11

suggests that our results would not differ significantly, especially in HY, if we were able to define size in terms of total amount outstanding. Multiple papers show that bond size is an effective "indirect" proxy for bond illiquidity (see Sarig and Warga, 1989 and Houweling et al., 2005).[42] Houweling et al. (2005) examine nine indirect liquidity measures, which, in addition to amount outstanding, include variables such as bond age and yield dispersion. They find eight of these measures to be effective (and highly correlated) measures of liquidity, with amount issued carrying the highest premium. Houweling and van Zundert (2017) note that this property of amount issued complicates the interpretation of a potential size premium. For our paper, the ability to isolate a size premium is of secondary importance to potentially capturing a liquidity premium as well.

We also adapt Downside risk and Credit quality factors from Bai et al. (2018). As these are novel factor definitions, they lack the extensive research of the previous factors. Bai et al. (2019) construct a Downside risk factor using indicators such as value-at-risk (VaR) and expected shortfall, which they suggest proxy for systematic risks, especially since bond returns are highly negatively skewed and exhibit high positive excess kurtosis. They intend for their Credit quality factor to capture systematic default risk, and find that credit ratings are just as effective as more advanced measures such as distance-to-default. Together with their Liquidity factor, they show that these factors explain a significant portion of the time-series and cross-sectional variation in corporate bond returns, and that Downside risk in particular carries a substantial risk premium. Additionally, Bali et al. (2019) borrow their four-factor model and show that it significantly outperforms traditional equity and bond factors in explaining the time series variation in corporate bond returns. Unfortunately, as they create their final factor portfolios using bi- and tri-variate sorts (as do Fama and French (2015)), our omission of the Liquidity factor means that our construction of the remaining factors will not align with theirs exactly. We do not see this as a problem, however, as their results do not change drastically from their initial univariate portfolio analysis, and because each factor exhibits individually strong explanatory power in the cross-section.

[42] The most effective measures of liquidity are typically classed as direct liquidity measures. According to Houweling et al. (2005), the highest order of such measures includes those based on transaction data, with variables such as the bid-ask spread falling one tier below.

3. ESG Literature Review

3.1 Overview of arguments for and against ESG-investing

The arguments for and against ESG integration in corporate bond portfolios come from the literature concerning ESG and portfolio construction broadly (Gerard, 2018). Scholars and investors who caution against the incorporation of socially responsible investment strategies suggest that ESG ratings are largely immaterial, and that portfolios sorted along such ratings will, on average, underperform due to diversification costs (Rudd, 1981). Alternatively, those in favor of ESG integration suggest that ESG ratings reflect relevant, extra-financial information, which help investors select securities that are more likely to outperform in the future. These two views are not necessarily mutually exclusive, but rather represent two opposing forces that influence the returns to any strategy involving investment screens. Regarding equities, the consensus is that, on average, the benefits to ESG integration outweigh any costs (related to the constraints ESG places on the investment opportunity set), such that ESG portfolios perform at least as well as conventional portfolios. The nascent literature on corporate bonds finds results vary depending on geographic context, but tends to agree that the performance of SRI funds is comparable to those of conventional funds (Leite and Cortez, 2016, Henke, 2016). There are reasons to suggest, however, that some of the arguments concerning the relationship between ESG and equity performance do not apply in the same way to bonds. These views also mirror the two main approaches towards ESG integration in portfolio construction. Some seek to optimize portfolios around ESG scores (Slimane et al., 2018) and others explore the extent to which active management can exploit the information contained in ESG ratings. The latter is the objective of this paper.

From the perspective of portfolio theory, the imposition of any constraints on the investment opportunity set translates into suboptimal investment decisions. The extent to which certain securities are excluded varies based on the ESG investment strategy employed, though each restricts the opportunity set to some extent. In regards to equity portfolios, Capelle et al. (2014) find that portfolios built using exclusionary social screens significantly underperform the market portfolio, while other types of screens do not have significant effects on performance. They also find that screening intensity is negatively related to financial performance, though an increasing number of screens partially offsets this initial negative effect, suggesting a potentially curvilinear effect. A number of researchers present methods to maximize a portfolio's ESG score while minimizing diversification costs and other negative effects on returns (Drut, 2010). For example, in the context of equity portfolios, Milevsky et al. (2006) present an optimization algorithm that effectively minimizes the costs related to the imposition of negative screens on passive index portfolios.

Some argue that the construction of such portfolios should be easier with fixed-income securities compared to equities, as bonds are more homogenous in regards to risk (sovereign and investment grade bonds in particular) (Hörter, 2016, suggests that ESG fixed income portfolios face fewer diversification costs). Drut (2010) finds that exclusionary screens can be implemented in sovereign bond portfolios without a significant loss of mean-variance efficiency, and suggests that asset managers can create sovereign bond portfolios with high average ESG ratings without significantly forgoing the potential for diversification. Li and Zhang (2017) find a similar result in regards to the application of sin screens on investment-grade, bond portfolios. They do not, however, examine the role of screening intensity and concede that stronger screens would eventually damage a bond portfolio's risk/reward profile. While this line of reasoning weakens the portfolio theory argument, it also diminishes the primary argument in favor of ESG investment strategies. If bond returns depend primarily on a few non-diversifiable risk factors, then there is also less room for active managers to exploit the information contained in ESG ratings. Derwall and Koedijk (2009) highlight that, especially in the high yield space, a significant portion of the risk of corporate bonds is firm specific and can therefore be significantly reduced through diversification or exploited by active management.

The argument in favor of ESG integration (that ESG ratings reflect relevant and potentially underappreciated information that portfolio managers can take advantage of in portfolio construction/security analysis) contains a number of components that each warrants separate investigation. The first concerns the relationship between ESG and economic performance at the firm level. Are these economic effects material for bond risk/pricing? Should ESG ratings reflect financially relevant information, the second major question is the extent to which this information is captured in traditional indicators, and, most importantly, in market prices. These analyses are crucial in order to distinguish persistent, economically interpretable patterns in average returns from temporary mispricing or statistical aberrations. For example, while ESG can influence asset prices in the absence of economic effects (due to investor preferences or expectations) these drivers tend to be less stable over time (with the exception, of course, of certain behavioral biases).

3.2 The economic effects of ESG

Scholars have debated the economic consequences of a firm's ESG activities in the literature for decades. Researchers have long called for the development of a unifying theory of corporate social performance (Ullman, 1985), but studies on the interrelationships between ESG reporting, ESG performance and economic performance yield inconsistent results. In one of the most broadly encompassing reviews of the literature, Margolis and Walsh (2003) identify 127 empirical studies on the CSP-CFP relationship published between 1972 and 2002. The majority of these studies (109) treat corporate social responsibility as an independent variable. Only four examine it as both an independent variable affecting financial performance and a dependent variable. The authors find that of these papers, 54 report a positive effect of CSR, 48 either report statistically insignificant results or otherwise find them to be inconclusive, and 7 provide evidence of a negative CSP-CFP relation. While other literature argues that the evidence is decidedly more in favour of CSR than many business scholars assume (Orlitzky et al. 2003, Adamska et al. 2016), empirical studies continue to provide evidence to the contrary. This highlights the need for more research.

This lack of consensus is partly attributable to a host of methodological issues that undermine the generalizability of past findings, including different procedures for data collection and analysis, inconsistent measurement of CSP and CFP, and variations in sample sizes as well as in geographic and industrial contexts (Liu and Taylor, 2019). Additionally, many researchers fail to address properly the issue that a firm's decision to invest in ESG is correlated with unobservable firm characteristics that also affect economic performance (an "endogeneity problem"). Similarly, a number of scholars suggest that the relationship between CSR and financial performance is bidirectional and simultaneous (Waddock and Graves, 1997, Orlitzky et al. 2003). As a result, most of the evidence presented in the literature does not clearly indicate a causal link between CSP and CSR in either direction. Also, given evidence that well performing firms are more likely to engage in CSR, some argue that the magnitude of the relationship is artificially distorted upwards (McWilliams and Siegel, 2000). Researchers have attempted to overcome this issue. For example, Edmans (2011), who uses lagged stock returns, has become the standard in many cases.

Among the most fundamental issues that continue to impede scholars' efforts to build a consensus, however, are disagreements regarding the definition and measurement of corporate social performance. This is particularly the case with terms like CSR and CSP. Some authors use CSR and CSP to refer to particular types of activities. Some activities concern exclusively environmental or social dimensions (Attig et al. 2013, Oikonomou et al. 2014), while others make more nuanced distinctions that depend on the environment in which a firm operates. For example, McWilliams et al. (2006) define CSR as pertaining to situations in which a firm engages in "actions that appear to further some social good, beyond the interests of the firm and that which is required by law" (McWilliams et al. 2006 pg. 1). Some studies, like that of Albuquerque et al. (2020), use E and S scores (in their case from Thomson Reuters), but exclude governance because they consider it "outside a firm's CSR remit". Others, such as Halling et al. (2020), do not include governance because they consider it priced (which some meta-studies, notably those from Deutsche Bank (2012) and UNPRI (2013), support).

Greater clarity and agreement regarding conceptual definitions of CSR must also accompany standards governing its measurement. Florian et al., (2020) find significant cross-rating variation in ESG ratings from providers that use similar definitions of ESG. The Sustainability Accounting Standards Board (SASB) is making great strides towards addressing these issues going forward, but it remains crucial to scrutinize the measures employed in the extant literature and to contextualize empirical results accordingly.

Ultimately, regardless how the literature defines them, terms like ESG and CSR inevitably encompass a wide range of activities, some of which may have different, and potentially opposing, effects on financial performance. A number of scholars assert that the use of broad measures of CSP has further obfuscated the true nature of CSP-CFP relationship (Brammer and Millington, 2008, Barnett and Salomon, 2012, Jayachandran et al. 2013), and recommend researchers analyze CSP at finer levels of disaggregation.[43] This paper reviews literature that focuses on the effects of specific environmental and social policies (King and Lenox, 2002, Edmans, 2011, Seltzer et al. 2019, to name a few), and also decomposes its ESG rating into various subcomponents in order to examine each individually in relation to bond returns. However, this paper contends that both its empirical findings and those in the existing literature suggest that it is best to view the economic effects of ESG/CSR within a broad framework.

The last major point of contention concerns the measurement of CFP. Some researchers argue that the use of short-term, accounting-based measures of profitability understate the effects of CSR on financial performance/firm value (Jiao, 2010, Edmans, 2011). This is largely a result of the contrast between CSR's immediate costs, which directly impact a firm's bottom-line, and its rewards, which are often largely intangible and realized over a longer time horizon (Oikonomou et al., 2014). For this reason Jiao (2010) advocates the use of Tobin's Q to measure the effects of CSR, since compared to measures of operating performance, Tobin's Q better captures the valuation effects of intangibles. It also helps researchers to avoid sample selection biases that arise from a focus on specific periods in which new information is released (Jiao, 2010). The influence of differing CFP measures is apparent in studies that examine similar issues using different indicators. For example, using traditional operating metrics, a number of researchers show that environmental controls, in particular pollution abatement, have increasing marginal costs, diminish productivity and ultimately lower economic performance (Gray and Shadbegian, 1993, Patten, 2002). Using Tobin's Q, Konar and Cohen (2001) and, to a lesser

[43] As most ESG ratings reflect many different metrics (KLD considers 304 metrics), even individual E, S and G pillars are so broad as to encompass activities with potentially contradicting effects.

extent, Guenster et al. (2011) arrive at the opposite conclusion, with the former finding that a 10% reduction in the release of toxic chemicals leads to an improvement in a firm's intangible asset value of more than $34 million. Increasingly, researchers use market prices, which as Edmans (2011) describes, both better captures the valuation effects of ESG and, when using returns, circumvents problems of endogeneity.

Lastly, many factors mediate the CSP-CFP relation. Some are exogenous, such as regional/country effects (Stellner et al., 2015), industrial context (Edmans, 2011), changes in regulation and policy uncertainty (Seltzer, 2020) and volatile public awareness/social pressure (Cappelle-Blanchard, 2012), while others relate to specific firm characteristics, such as insider ownership (Bhoraj Sengupta, 2003) or firm size (Stanwick and Stanwick). Adamska et al. (2016) suggest that the lack of research on the effects of particular firm characteristics that impact the CSP-CFP relationship undermines evidence of a positive association. Specifically, they argue that the evidence of positive returns to CSR investment have been of a specific, rather than universal nature, and that the literature gives little insight into which factors determine the effects of CSR.

3.2.1 Two major views regarding ESG-CFP relationship

In the midst of mixed evidence, two largely dichotomous views concerning the CSP-CFP relationship have survived alongside one another (Adamska et al. 2016). Liu and Taylor (2019) refer to these two views as the "traditionalist" and "revisionist" perspectives, the definitions of which this paper will borrow, and build upon, in order to summarize the literature concerning the economic effects of ESG. Put simply, the traditionalist view assumes that incremental spending on ESG destroys shareholder value, while the revisionist view asserts that firms engage in profit-maximizing ESG (Albuquerque et al. 2019). When assessing specific environmental or social policies, traditionalists tend to emphasize the immediate demands such policies place on a company's scarce resources (Barber, 2007). Revisionists, on the other hand, identify a range of advantages to CSR that translate into superior financial performance in the long run, including product differentiation (Albuquerque et al. 2019), decreased regulatory/litigation risk (Koh et al. 2014), reputational gains (Edmans, 2011) and improved innovation and/or productivity (Hart, 1995, Porter, 1995). What makes these perspectives so difficult to reconcile, however, is that they reflect a larger rift within modern economics. The neo-classical view of the firm stresses the primacy of shareholder interests, and non-traditional behavioural theories of the firm assert that companies consist of many different individuals and groups whose time-varying (and often conflicting) needs require managers to maintain multi-dimensional objectives (Menz, 2010).

3.2.2 The traditionalist view

The conceptual framework underlying traditionalist arguments against CSR comes from agency theory. Introduced in the seminal work of Jensen and Meckling (1976), agency theory highlights the considerable risks to which shareholders are exposed given that managers can, and often do, make decisions out of self-interest or incompetency that run counter to the objective of shareholder value maximization. The authors even list donations to charitable organizations and close employee-management relations as activities that pose high risks of managerial malfeasance. To the extent that classical finance theorists remain unwavering in their view that CSR/ESG reflects a costly diversion of scarce internal resources, traditionalists argue that corporate sustainability is itself a type of agency cost, such that managers accrue private benefits from embedding environmental or social policies into a company's strategy (Eccles, 2013). Friedman most famously articulated this view in his 1970 article, in which he argued against growing support among business leaders for the notion that firms bear a responsibility to help ameliorate societal ills. As the father of shareholder theory, Friedman asserted that a firm's efforts to maximize shareholder wealth, while occasionally resulting in negative externalities for other stakeholder groups, ultimately maximizes societal wealth as well. To deal with such externalities is the domain of governments, he argued, and for companies to address such issues amounts to managers spending shareholder money to pursue personal social, political and economic goals (Friedman, 1970).

Although not as popular as it was prior to the new millennia, the view that agency problems may manifest as managerial altruism or social preferences (Fehr and Schimdt, 1999; Charness and Rabin, 2002), which have negative ramifications for firm value (Brown et al., 2006) continues to find support in the literature. Barnea and Rubin (2005) highlight the increasing desire of managers to appear as industry leaders in regards to sustainability, and equate CSR investments to purchases of unnecessary corporate jets (Yermack, 2006). Others attribute more sinister motives to managers, suggesting that they use CSR as a means of concealing corporate misbehaviour in conjunction with shoddy accounting practices (Kim et al. 2014). One of the better-supported arguments concerns CSR as a form of managerial entrenchment. It is here that traditionalists agree with revisionists that managers seek to promote stakeholder welfare through CSR. Unlike revisionists, however, traditionalists view such engagement as a means of resisting shareholder discipline (Zwiebel 1995). Although it is difficult to discern the extent to which CSR reflects managers' efforts to protect their interests that entrenchment is detrimental to shareholder value is well established. Bebchuk et al. (2009) document a statistically and economically significant negative relationship between an entrenchment index and firm value, while Pagano and Volpin (2005) show that incumbent managers exploit pleasant labour-management relations to deter value-adding takeover bids.

Lastly, the view that CSR represents a type of agency cost overlaps with another traditionalist argument commonly referred to as the "overinvestment view" (Barth et al. 2018). Although softer in its assessment of both managerial incentives and the benefits to CSR, it highlights CSR's increasing marginal costs and suggests that, due to either social pressure and/or an inability to adequately value CSR opportunities, companies overspend on "corporate goodness", thus diverting valuable resources from higher value creating opportunities (Barber, 2007).

3.2.3 The revisionist view

One of the primary faults revisionists find with traditionalist arguments is the implicit assumption that a corporation cannot act for the benefit (and profit) of others without neglecting its shareholders (Attig et al. 2013). In an attempt to reconcile CSR with the shareholder view of the firm, revisionists suggest that firms' efforts to appease multiple constituencies need not be detrimental to shareholder interests. To the contrary, revisionists argue that doing so improves financial performance. This is the basis of stakeholder theory, which Freeman (1984) developed in opposition to Friedman's shareholder theory and which revisionists use as their theoretical foundation. According to stakeholder theory, shareholders are but one of many interrelated stakeholder groups which collectively comprise the corporate environment. Freeman defines stakeholders as groups "without whose support the organization would cease to exist" (Freeman 1984, pg. 13), including employees, customers, suppliers and shareholders (what scholars typically refer to as primary stakeholder groups), as well as local communities, governmental groups and the environment. Stakeholder theorists argue that, to succeed over the long term, a company must consider and satisfy the needs of all stakeholder groups.

At the most general level, revisionists use stakeholder theory to advance "doing well by doing good" arguments (Jiao, 2010), such that companies exhibiting high levels of social responsibility and sustainability are more likely to succeed in all dimensions, including financial performance. Extensive research supports the vital role stakeholders play in firm performance (Porter and Kramer, 1997, Eccles, 2013, Salvi et al., 2019), with some researchers suggesting that stakeholder welfare be considered alongside the likes of human capital and reputation as intangibles that are critical to a firm's success (Jiao, 2010). Importantly, stakeholder theorists not only emphasize the benefits firms accrue by addressing the needs of their stakeholders, but also the risks they face should they fail to do so (Eccles et al. 2013). Examples of the former include improved worker morale and productivity (Edmans, 2011), enhanced customer loyalty and product differentiation (Albuquerque et al, 2018). Examples of the latter include difficulty hiring talented workers (Greening and Turban, 2000), increased fines due to litigation and regulation (Bauer and Hann, 2010) and consumer boycotts (Sen et al. 2001). The literature also provides some evidence that the market recognizes the value of stakeholder welfare and CSR's role in promoting it. Hillman and Keim (2001) find a positive relationship between an aggregate CSR measure and CFP for a sample that includes the constituents of the S&P 500. Crucially, they find that scores related to primary stakeholder welfare are the sole drivers of this relationship, while other dimensions embedded in CSR scores exhibit a negative relationship with CFP. As a result, they conclude that primary stakeholders are much more important than social issue participation for shareholder value creation. Similarly, from a sample of 822 U.S. based firms in the period from 1992 to 2003, Jiao (2010) shows that various dimensions of stakeholder welfare, in particular welfare scores related to employee relations and environmental performance, are positively and significantly related to Tobin's Q. Jiao (2010) argues that the valuation effects associated with stakeholder welfare support the view that stakeholders represent valuable intangibles like human capital and reputation, and directly opposes the traditionalist view that managers' pursuit of private benefits drive stakeholder welfare.

Revisionists derive from this basic-framework, two-core argument in favor of CSR, which scholars often refer to as the "good management" and the "good company" hypotheses (Gerard, 2018). The "good management" hypothesis suggests that CSR signals managerial quality, so much so that some researchers and investors use ESG scores as proxies for management strength (Bucholz 1978, Bollen 2007). Waddock and Graves (1997) argue that to effectively integrate ESG into a company's strategy requires management to navigate the diverse interests of multiple stakeholder groups and to balance tangible costs against intangible rewards, both of which are essential for companies to maximize long-term value (Edmans, 2011). Additionally, theories of managerial myopia contend that managers tend to underinvest in intangibles because the market only values them in the very long run (Stein, 1988). High levels of CSR, therefore, indicate management's ability to overcome the problems associated with corporate myopia (Edmans, 2009). Lastly, in the context of strategic management theory, Waldman et al. (2004) suggest that certain aspects of transformational leadership relate to positively to CSR, supporting the view that strong managers recognize the strategic value of CSR.

The relationship between various characteristics of capable managers and a company's corporate social performance has been demonstrated empirically. Most recently, Eccles et al. (2013) find for a large cross-section of U.S. companies that, compared to low sustainability firms, high sustainability firms (based on KLD scores) more effectively identify, and engage with, key stakeholder groups and are more long-term oriented. The authors also show that such firms outperform their low CSR peers using both accounting and market-based metrics. There is also moderately strong empirical evidence that high CSR firms face fewer agency frictions and a lower likelihood of short-term opportunistic behavior by managers (Benabou and Tirole 2010), both of which directly undermine the traditionalist view of CSR. Using a sample of more than 7,000 companies worldwide over the 1999-2011 period, Ferrell et al. (2016) provide one of the most comprehensive examinations of the link between CSR, agency problems and firm value. The authors find that companies that exhibit lower agency costs engage in more CSR, and that the positive relationship between CSR and firm value is more pronounced among low agency problem firms. They also document that CSR activities relate positively with executive pay-performance sensitivity, a finding that is difficult to reconcile with the view that CSR is motivated by managers' personal agendas. Additionally, Gao et al. (2014) show that executives are less likely to engage in insider trading at high-CSR firms compared to those at low CSR firms, while Kim et al. (2012) find that high-CSR firms are less likely to manage earnings.

The "good company" hypothesis emphasizes the broader significance of a company's track record in addressing the interests of its stakeholders. It asserts that stakeholder engagement via ESG generates reputational capital that improves firm value through enhanced profitability and/or protection against adverse shocks (Hillman and Keim, 2001). This view can be further distilled into two related components, one which emphasizes the benefits of CSR as a marketing strategy (Baron, 2001) and another that conceptualizes trust and cooperation as strategic assets (Jones, 1995). Regarding the latter, evidence supports the value of ESG in product differentiation (Klein and Dawar 2004, Brown and Dacin 1997, Berens et al. 2005), with some customers exhibiting a willingness to pay a premium for sustainable products (Sen & Bhattacharya 2001). Furthermore, Turban & Greening (1997) document a positive relationship between a firm's corporate social performance and its external reputation as an employer, and a number of scholars suggest that high sustainability firms are able to attract qualified employees at lower costs (Phillips et al. 2007, Albuquerque et al. 2018).

In addition to strong branding, CSR helps firms to build fruitful relationships with its stakeholders. Jones (1995) highlights the value of social capital, which a firm accumulates should its stakeholders perceive it to be a "cooperative and trustworthy partner". For example, employees who feel they have been fairly treated and remunerated exhibit higher working morale and stronger commitments to the success of the company (Fehr and Falk 1999). In the most thorough examination of the interrelationships between a firm's CSP record, employee satisfaction and firm value, Edmans (2011) finds that a firm's track record in building a supportive and cooperative corporate working environment promotes both job embeddedness (which ensures talented employees remain with the firm) and worker morale (which improves worker productivity).

Building on Jones (1995), Godfrey et al. (2009) highlights the importance of maintaining positive relationships with society at large, and argues that social capital (or rather his related concept of "moral capital") has insurance-like qualities that positively affects valuation. Chatterji et al (2009) demonstrate that firms with poor social performance face significantly more pollution and regulatory compliance violations than other firms. Similarly, Koh et al. (2014) find that various ESG activities (including sincere corporate governance, employee satisfaction and strong environmental performance) significantly decrease future litigation risks. In addition to less frequent litigation, high CSR firms also appear to benefit from lower costs in cases where they commit violations. Using prosecutions of the Foreign Corrupt Practices Act, Hong and Liskovich (2016) report that more socially responsible firms pay lower fines for bribery when violating the Foreign Corrupt Practices Act. Similarly, Jeffers (2015) finds that officials are more lenient with penalties for OSHA violations ascribed to high CSR firms. Lins et al. (2017) find that high ESG firms perform best during a crisis in terms of accounting and market-based measures. They attribute this result to social capital, which they argue engenders reciprocity (stakeholders, including shareholders, are more likely to help a firm with high social capital during a shock due to trust).

3.2.4 Relevance of ESG Risks Over Time

ESG risks may have become more material over time. The relationship between ESG and bond returns over time is important to determine whether the ESG effect supports a demand or risk-based explanation. That ESG may have become material in recent years could obscure either effect in the time series. The role of CSR in reducing firm risk is arguably becoming more important as physical environmental risk, together with regulation and public scrutiny regarding sustainability issues, continue to rise. Henke (2016) notes the rising materiality of ESG risks over the past few decades. He explains that climate-related events resulting in severe losses have more than doubled from fewer than 400 events per year in the 1980s to more than 800 events per year after 2000, while the financial crisis highlighted the scale of corporate governance risks (financial institutions have paid more than 275bn USD in fines). Regarding social risks, Henke (2016) references the consumer outcry and product boycotts that followed the 2013 collapse of the Rana Plaza factory building in Bangladesh. Holt and Barkmeyer (2012) show that rising public awareness concerning issues such as worker conditions and safety channels, waste management and the risk of consumer boycotts as a result of sustainability issues has risen substantially (cite). Jagannathan et al. (2017) show that the risks stemming from sudden changes in consumer preferences or regulatory changes, while rare, can cause such large asset price swings that optimal portfolio construction should explicitly account for such risks. Lastly, Godfrey et al. (2009), Goss and Roberts (2011), El Ghoul et al. (2011) and Bouslah et al. (2013**)** show that CSR results in a significant decrease in idiosyncratic firm risk. Alternatively, Albuquerque et al. (2020) develop a theoretical model in which, through investments into CSR, a firm's efforts to differentiate its products reduces its exposure to systemic risks and increases firm value.

Although many of the conceptual and methodological issues outlined above make it difficult to discern the exact effect of CSR on financial performance, in particular due the range and variability in corporate activities that fall under the CSR umbrella, the empirical evidence is very difficult to reconcile with the traditionalist view of sustainability. On the other hand, there is ample evidence that firms recognize the strategic value of CSR, and, especially given the rise in environmental and social risks, this paper contends that the revisionist framework is a superior guide in terms of the development of a theoretical framework for the relationship between CSP-CFP.

3.2.5 The Moderating Influence of Agency Risks and Governance on ESG-CFP relationship

Most of literature on the ESG-CFP relation seems to argue either for or against ESG. We prefer to highlight the few papers that arrive at more nuanced conclusions. These include 1) potential trade-off between CSR and governance and 2) that factors such as region or industry may determine whether ESG adds or detracts. Cheng et al. (2020) attribute a significant amount of CSR spending to agency problems. The authors examine firm behavior before and after the 2003 dividend tax cut in the United States, an event that, in many cases, accompanied an abrupt increase in insider ownership. The authors assert that this spike in insider ownership accompanied a decrease in agency costs as insiders rechanneled unproductive spending to more-productive investments or to shareholders via dividend payouts (there was a sharp increase in dividend payouts following the tax cut). They find that firms that an increase in insider ownership following the tax cut was associated with an increase in firm value and a decrease in CSR spending. They therefore attribute a firm's marginal spending on CSR to agency problems, which they argue has negative financial ramifications. Albuquerque et al. (2019) argue against the results of a very similar paper (Masulis and Reza, 2015, 2019), noting that while they show how a change in after-tax ownership impacts CSR, they do not show that CSR activities destroy value.

Still, the results of Cheng et al. (2020) imply a potential trade-off between CSR (E and S) and governance, something that previous papers have, to varying degrees, supported. Notably, Krüger (2015) finds that, in situations where managers are likely to receive private benefits from the adoption of CSR (which could apply to cases where there are weak governance mechanisms in place), that the marginal effects of CSR spending are negative. However, Krüger shows that CSR has a positive effect on firm value if that firm adopts CSR policies to improve relations with stakeholders.

Other factors that might influence whether ESG adds or detracts value include the region or industry in which a firm operates. Breuer et al (2018) document lower costs of equity for high ESG firms only in countries where investor protection is high. The idea is that insufficient investor protection could enable corporate managers to extract private benefits through ESG investments. They find that investor protection in Europe is higher than in the U.S., which they argue is consistent with lower costs of equity for European firms. Barth et al. (2019) also find a great risk-mitigation effect of ESG in Europe. Menz (2010) and Stellner et al. (2015), however, arrive at opposing results.

3.3 ESG in fixed income

The literature identifies the cost of capital as another major channel through which ESG influences financial performance. As Sharfman and Fernando (2008) note, perceived changes in a firm's riskiness due to ESG may lead to reduced costs of capital, which, commensurate with a decrease in a firm's overall cost base, increases its ability to generate profit from a given level of revenue. Additionally, measures of a company's cost of debt financing, such as its credit rating or credit spread, are not only both indicators of a company's financial health and strong predictors of its future financial performance, but also circumvent many of the problems associated with short-term, accounting-based measures of profitability (insert footnote). As a result, many of the researchers who examine the relationship between ESG and the cost of debt have done so in an attempt to uncover the "unique and thus far unrevealed roles" (Sun and Cui 2014 pg. 4) of CSR in regards to CFP, rather than to understand ESG's implications for bond pricing.

3.3.1 Effects and implications of ESG on bond pricing

While the literature concerning the economic effects of CSR/ESG stretches back a half century, it is only relatively recently that researchers have examined the relationship between ESG and debt. Of the 52 and 103 studies reviewed by Ortlizky et al. (2003), and Margolis and Walsh (2003) respectively, not one explicitly examines the link between CSR and the cost of debt (Goss and Roberts, 2011). This is somewhat surprising given the size and significance of the corporate bond market, but it is consistent with the general paucity of research on corporate bonds relative equities (due to, among other factors, limited data and the complex nature of bond pricing (Gerard, 2018). Not only is debt the "marginal source" of financing for many companies, (Gourio (2013)), but the corporate credit spread is an excellent predictor of future economic activity. Over the past two decades, researchers have increasingly turned their attention to the relationship between ESG and key bond metrics, notably credit ratings and the cost of debt financing. Some of these scholars appear more interested in examining credit/default risk in order to illuminate the "unique and thus far unrevealed roles" (Sun and Cui, 2014, pg.4) of ESG in terms a firm's financial performance. Many of the papers in Deutsche Bank's (2012) meta-study fall under this category. They also expand their scope to include research that examines the cost of capital more broadly. Recent papers, however, focus more on the role of ESG in bond pricing.

There is substantial evidence of an inverse relationship between ESG concerns and the cost of debt, and while the evidence with respect to ESG strengths is less convincing, it is more supportive of a negative or neutral relationship than a positive one. On the one hand, it is very encouraging that these results are consistent across a wide range of definitions and measures concerning ESG as well as different indicators of the cost of debt. On the other hand, the variation in these measures raises many of the same methodological and data-related issues that continue to obfuscate the relationship between ESG and financial performance more broadly. Notwithstanding the many ESG/CSR ratings from different providers, there is significant variation in researchers' measures of credit risk, as well as in their definitions of corporate sustainability. Regarding the former, most studies either examine indicators related to tradeable debt, such as bond spreads in the primary (Amirslani et al. 2019, Halling et al. 2020) and secondary markets, credit ratings (Jiraporn, Sun and Cui (2014), Yang (2020), and CDS spreads (Barth et al. 2020), or to non-tradeable debt, such as rates on bank loans (Goss and Roberts, 2011, Chava, 2014) and estimates of debt-related costs (Magnelli, 2017). There are also some who investigate the relationship between ESG and factors that influence the cost of debt, such as bond covenants (Shi and Shun 2015), access to financing (Cheng 2014) and a company's financing policy (Sharfman and Fernando, 2008, Bali et al. 2009). In terms of corporate sustainability, some papers examine aggregate CSR or ESG scores (Attig et al. 2013, Oikonomou et al. 2014, Salvi et al. 2019), while others focus on environmental (Bauer and Hann, 2014, Chava, 2014), social (Edmans, 2011) and governance (Ashbaugh, 2006) pillars separately.

These different measures of the cost of debt reflect slightly different information. Spreads on tradeable and non-tradeable debt, for example, reflect not only remuneration for credit risk, but also a systemic risk component, and, in the case of tradeable debt, liquidity risk. These three factors do not completely explain the determinants of credit spreads, however. Additionally, there is a component of investor/lender preferences embedded in bond/loan spreads, which, although rating agencies are not unbiased (Chava et al., 2014), are not reflected in the same way into credit ratings.

The theoretical channels through which ESG may influence credit spreads or bond returns requires further investigation. As Halling et al. (2020) illustrate, a simple model could decompose credit spreads into three separate channels. These channels would illustrate the factors that influence the level of cash flow (risk-based), the stability of cash flow (risk-based) and investor preferences (demand-based). Merton (1974) links the economic arguments to credit. Should ESG result in higher and more stable cash flows, holding fixed all other factors that influence bond prices, this should translate into higher asset values and thus a lower probability of default and lower credit spreads. Alternatively, ESG can reduce volatility in firm value by reducing idiosyncratic risk thus stabilizing its competitive position. This similarly reduces probabilities of default and thus lowers bond spreads.

3.3.2 Contradictory results

Only three papers document a positive relationship between ESG and the cost of debt, and in each case there are substantial factors that may explain their results. For example, while Scharfman (2008) finds that, on average, firms with strong environmental performance incur a greater cost of debt, he does not control for the impact of a firm's total degree of leverage. He documents that a firm's environmental performance is also positively related to its level of debt, which is consistent with Cheng (2014) and Amirslani (2019) who find that ESG improves access to financing, particularly with regard to debt markets. Scharfman (2008) concedes that the relationship between his measure of environmental performance and a firm's level of debt may explain his finding of a positive relationship between environmental performance and spreads.

Menz (2010) and Magnanelli and Izzo (2017) report a positive relationship between CSR and the cost of debt even after accounting for a range of borrower and bond characteristics that previous studies have identified as cross-sectional determinants of bond yields/credit risk. However, both papers use the same measure of CSP (DJSI) and examine European bonds (or at least a Euro-heavy sample in the case of Magnanelli and Izzo (2017) over a similar period.

In a sample that covers 498 European corporate bonds over the period of 2004 to 2007, Menz (2010) finds that the bond yields of socially responsible companies were *ceteris paribus* higher for than non-socially responsible companies. This result, however, is only marginally significant, and Menz (2010) suggests that, rather than indicating a true positive relationship between CSR and the cost of debt, that it reflects that CSR has yet to be incorporated into the pricing of European corporate bonds. In support of this interpretation, later papers contend that investors in Europe have been slower to incorporate ESG information into stock prices compared to investors in the United States (Badia and Cortez, 2020). Menz (2010) also uses an uncommon measure of CSR (SAM ratings), and he suggests that, at least from the perspective of bondholders, this indicator may not be a good proxy for firms' CSR (although consistent with Badia and Cortez (2020), Slimane (2018), contradicts this and claims Europe is where ESG in fixed income most advanced.

Magnanelli and Izzo (2017) document a positive relationship between an uncommon measure of CSP (borrowed from the Dow Jones Sustainability Index rating) and an uncommon indicator of the cost of debt (measured as the ratio of "financial interests expenses on financial debt," which is a proxy for the cost of debt provided by Thomson Reuters). Additionally, their sample size is small considering it comprises of companies across countries, industries and ratings classes and it only covers four years. Stellner et al. (2015) also corroborate a weak connection between aggregated ESG ratings and corporate bonds in terms of z spreads.

3.4 Further Evidence in Fixed Income

For a large sample of U.S. based firms from 2003 to 2011, Attig et al. (2013) document a significant positive relationship between a firm's CSR score and its credit rating, a result that remains robust to a range of alternative sample specifications and additional controls (including endogeneity). Importantly, they show that the dimensions of their CSR score that drive this relationship are those related to a firm's primary stakeholders, such as workforce diversity, employee relations and issues concerning products, the environment and the community (Attig et al. 2013, pg. 686). Alternatively, social issues not directly related to a firm's primary stakeholders, notably human rights, exhibit no relationship with credit ratings. Upon further inspection of the constituent CSR measures, the authors argue that strengths indicate a company's commitment to CSR that exceeds its direct economic interests. Therefore, consistent with the revisionist concept of profit-maximizing CSR, they assert that their results reflect discretionary (as opposed to altruistic) behavior of firms, which decide to delegate scarce resources to address stakeholders' interests with the purpose of maximizing their competitive position.

Oikonomou et al. (2014) examine a comparable sample of U.S. based firms, but use different ESG ratings and consider the influence of CSR on corporate bond spreads in addition to credit ratings. Like Attig et al. (2013), they decompose their aggregate CSR measure into qualitative business areas that they tie to particular stakeholder groups, including employee relations, product quality and issues related to the community and the environment. For both the aggregate measure and the dimensional components, they separate CSP strengths from CSP concerns. The authors show that, in general, the bond market rewards good CSP performance in the form of lower yields and stronger ratings, while it penalizes social and environmental transgressions with higher yields and weaker credit ratings. Like Attig et al. (2013) and Oikonomou et al. (2014), results suggest that the bond market recognizes the value of stakeholder welfare, but their separation of strengths from concerns allows them to provide a more nuanced analysis of specific welfare scores. For example, although they find that an avoidance of employee controversies significantly increases credit risk (all else equal, employee controversies can lead up to an 80% increase in a firm's credit spread), they do not find evidence that firms with strong employee relations benefit from lower spreads/higher credit ratings. The opposite is true of higher levels of product safety and quality as well as support for local communities. Lastly, in contrast to Attig et al. (2013), they find a beneficial effect of diversity issues on debt financing. Although this relationship is the least significant from an economic perspective, that they are unable to reconcile this result with past research. The revisionist framework highlights the impact of methodological and data issues on the precision of empirical analysis on CSR.

Salvi et al. (2019) provide the most recent, and the most comprehensive, examination of the impact of specific stakeholder dynamics on credit quality and debt pricing. Using one of the largest, international samples of bonds in the CSP-debt literature, the authors affirm the general results of Attig et al. (2014) and Oikonomou et al. (2014) (such that issuers with more CSP-related controversies face weaker credit ratings and higher costs of corporate bonds, while those with CSP strengths benefit from better ratings and lower credit spreads). Aside from the size of their sample, however, what separates Salvi's paper is their construction of stakeholder indices. They identify six groups that they argue have the greatest connection to a firm's survival and long-term license to operate. These include customers, employees (which they combine with suppliers), the community, shareholders, a firm's management and the environment.[44] Although Salvi et al. find that all stakeholder groups are significantly related to credit rating decisions (such that each important stakeholder group has the capacity to cause a hit or a rise in a firm's credit quality profile), not all groups appear to impact bondholders' perception of the issuing company's risk. In particular, practices related to employees, management, customers and the community exhibit economically and statistically significant negative relationships to the cost of bonds. In contrast, they find that practices linked to shareholders and the environment do not appear to have a significant influence on spreads.

These results are interesting for at least three reasons, some of which this paper will discuss in greater detail. First, the different results for credit ratings versus credit spreads raises questions regarding the marginal information CSP scores provide after controlling for credit ratings. The more credit ratings subsume the information contained in ESG scores, the less active managers will be able to exploit ESG ratings. Second, while their results concerning employees and customers are consistent with some of the better established parts of the revisionist literature, their results with respect to their

[44] They acknowledge that the community, the environment and management are not typically considered primary stakeholders (with the latter rarely conceived as a separate stakeholder group), but argue that doing so is justified given the evolution of the corporate environment due to CSR and stakeholder relations.

management and community scores are relatively novel. This is partly because it is very rare for researchers to consider management as a distinct stakeholder group. The components of the management score that appear to drive the negative relationship with credit risk are those measuring managerial alignment with the firm's objectives, compensation alignment and the level of ESG integration within the firm's overall strategy, all of which suggest that the bond market appreciates the value of "good management" reasoning. Additionally, the authors argue that the community score measures a firm's philanthropic attitude. They therefore attribute the negative relationship between the community score and credit spreads to the bond market's appreciation of firms that assume a good citizen role.

Evidence that bondholders recognize both tangible and intangible rewards to strong CSR makes it even more surprising that Salvi et al. (2019) fail to detect any relationship between their environment and shareholder scores. They attribute this result to the mediating role of regulation. They examine the role of a country's legal context on the CSR-debt relationship, and find that the extent to which a country's legal system emphasizes or explicitly regulates particular stakeholder relations impacts the influence such relations have on the cost of debt. They find that in common-law countries where most of the activities subsumed under the shareholder and environment score are highly regulated, the scores exhibit no relationship to credit spreads, whereas in civil law countries in which other aspects are better regulated, the relationship becomes negative. Salvi et al. (2019) suggest that the regulation of environmental practices and corporate governance suppresses the relationship between bonds and the environment/shareholders. Elsewhere in the literature, authors explore the idea that there are different rewards to voluntary versus proactive CSR but this theme is less relevant for this paper.

Halling et al. (2021) examine the relationship between environmental and social performance and credit spreads in the primary bond market. They contend that primary markets represent a cleaner setting in which to examine the effects of ESG, as offering prices are less likely to be influenced by temporary changes in liquidity levels and, at the issue stage, bonds carry a recent credit rating, which effectively controls for many issuer and bond characteristics. For a large sample of U.S. issues from 2002 to 2020, the authors find that the primary bond market rewards firms with good ESG performance, even after controlling for bond ratings and various firm characteristics. They find, however, substantial variation across the different components of their ES-score. They find that product-related components, and to a lesser extent employee relations, account for the majority of the observed relationship. Dimensions related to diversity, the environment and human rights (which get significantly more attention in the media and from social activists), exhibit no statistically significant relationship with issue spreads in the primary bond market.

For a large cross-section of US public companies, Bauer and Hann (2014) document that environmental concerns are associated with a higher cost of debt financing, while environmental strengths are related to a lower cost of debt (Bauer and Hann 2014). They find that, holding all other variables fixed, the gap between firms with strong environmental performance versus those with weak performance is as large as 53bps annually. Among the most important environmental strengths that explain the observed negative relation with bond spreads, the authors identify a firm's commitment to climate-friendly policies and practices, or its efforts to reduce air pollution and its impact on climate change broadly (also supply of innovative products with environmental benefits). Importantly, they stress the role of environmental practices in mitigating risk as opposed to generating revenue. They note that environmental liabilities, as well as consumer and employee safety or equal opportunity cases, can result in risks so significant as to tempt management to consider filing for strategic bankruptcy in an effort to avoid them.

They build their conceptual framework around the view that environmental practices influence the solvency of borrowing firms by determining their exposure to legal, reputational and regulatory risks. Both penalties for environmental misconduct, as well as strong negative reactions from a firm's various stakeholders, have the capacity to impact a firm's default risk and therefore to impair the value of their fixed income securities. They also expect that this relation will strengthen over time as widespread climate change concerns have heightened investors' awareness of potential regulatory changes and other associated financial risks.

3.5 (Mis)alignment of bondholder and shareholder interests

The results discussed so far have been consistent with the literature concerning the relationship between CSR and firm value. However, corporate bond spreads do not necessarily mirror changes in firm value. Although equities and bonds represent claims against the same underlying assets (Merton, 1974), the way they respond to changes in asset values is not the same (Israel et al., 2018). Additionally, bond values can change even if the fundamental value of the underlying business does not. This is partially a result of the potentially divergent interests of bondholders and shareholders due to differences in terms of capital structure. Shareholders are owners of the firm, their payoffs are unknown and stochastic and they participate (equally in the upside as they do the downside: rephrase). As the ultimate "fixed claimants", bondholders are most concerned with the risk that a firm might default before bond maturity, and thus their primary objective is to manage default risk. Since they have no (or at least comparatively limited) variable upside to compensate them for this risk, they typically have a lower risk tolerance than do shareholders. This leads to a number of potential conflicts between shareholders and bondholders, which are broadly referred to as the "agency costs of debt."

In many cases, managers face incentives to act in ways that benefit shareholders at the expense of bondholders. For example, managers of financially distressed firms may have an incentive to pay out cash to shareholders in the form of dividends or stock repurchases prior to bankruptcy, effectively diverting cash from bondholders to shareholders (Wald and Long, 2007). Alternatively, shareholders of such firms may pressure managers to invest in excessively risky projects, even if such projects would likely to reduce firm value, since they provide an opportunity for shareholders to receive a much higher payoff (Williamson, 1988), Amiraslani et al. 2019). There may also be situations where the interests of bondholders conflict with the maximization of firm value. As Jensen and Meckling (1976) explain, companies must take risks to succeed long-term, but bondholders, who as a function of their pay-off structure are typically interested in a safer investment, may want to place restrictions on managers' use of money to reduce risk, even if doing so may negatively impact value in the long run. Lastly, disciplinary takeovers (in particular leveraged buyouts) often benefit target shareholders, but hurt bondholders as the firm's debt level increases (Cremars et al. 2007). In all cases, debt agency costs reflect situations where bond value may change irrespective of the value of the underlying corporation.

In the context of debt agency costs, bondholders and shareholders appear to perceive certain governance mechanisms differently. This is particularly the case regarding board independence, the concentration of institutional ownership and antitakeover provisions. Gompers et al. (2003) were the first in a string of papers to analyze the relationship between their "G-index" (a measure of anti-takeover provisions that they used as a proxy for shareholder rights) and firm value. They argue that strong anti-takeover provisions reflect greater managerial entrenchment and ultimately lead to the divergence of managers' interests from those of shareholders. They document that firms with strong shareholder rights (and therefore weak antitakeover provisions) exhibit positive financial performance (due largely to lower capital expenditures and fewer acquisitions) and earn positive abnormal long-run stock returns. Klock et al. (2005) find the opposite result regarding a similar governance index and the cost of debt financing, such that, all else equal, firms with strong antitakeover provisions benefit from a lower cost of debt financing. Ashbaugh-Skaife et al. (2006) and Cremars (2007) find similar results (in terms of credit ratings and bond yields respectively), which suggests that stronger management rights, although not beneficial to shareholders, are viewed favorably by the bond market.

Cremars (2007) shows that takeover vulnerability mediates the relationship between factors concerning shareholder control and bond yields. The difference in bond yields between firms with strong and weak shareholder control grows almost linearly with takeover vulnerability, up to a maximum difference of 67bps, all else equal. Additionally, he finds that bond covenants diminish the credit risks associated with strong shareholder control, and highlights the role such covenants play in ensuring that bondholder and shareholder interests do not diverge. Chen et al. (2007) provide further evidence that bondholders are willing to pay a premium to minimize shareholder control, with unionized firms facing significantly lower costs of debt than non-unionized firms all else equal (unions mitigate the tendency for shareholders to expropriate bondholders).

A number of scholars suggest that, given the agency costs of debt, the “good management” and “good company” reasoning regarding CSR is of particular importance to bondholders. Shi, Sun and Su (2015) argue that should ESG indicate management’s capacity to balance the interests of different stakeholders and to minimize agency frictions, this should be reflected in the number (or “intensity”) of bond covenants. They find that strong CSR is associated with fewer bond covenants, while CSR concerns are related to a higher number of covenants. They replicate this result across four main covenant types, including event covenants, investment covenants, dividend covenants and subsequent financing covenants. Their results are most pronounced with the latter two, which is consistent with findings in the literature that suggest stockholders are most likely to use dividend and financing policy to “hurt” bondholders (Smith and Warner, 1979). Additionally, they find that firms with high bid-ask spreads and high agency costs show a stronger relationship between CSR and bond covenants. They attribute their results to CSR’s impact on bondholders’ perception of risks, firm reputation as well as a reduction in information asymmetry due to the extra-financial information ESG ratings provide.

Amiraslani et al. (2019) present the agency costs of debt as the primary channel through which ESG influences bond yields. Building off the “good company” hypothesis, they suggest that ESG scores reflect the trust a firm has accumulated with its stakeholders. They argue that the most appropriate setting in which examine the value of social capital is during a system-wide shock to trust, and they therefore hone in on the 2007-2008 financial crisis. Since the crisis is a plausibly exogenous event regarding firms’ decisions to invest in CSR, this strategy also allows them to bypass some of the endogeneity problems that plague most studies concerning the capital market effects of CSR. Using an aggregate measure of CSR as a proxy for social capital, they find only a moderately negative CSR-credit spread relation for their full sample (include time period and size), which disappears entirely after they control for time fixed effects. During the crisis, however, they find strong evidence that bond spreads of high-CSR firms did not rise as much as the spreads of low-CSR firms. Moreover, they find that firms with greater opportunities to engage in asset substitution or cash diversion (non-investment grade firms with fewer intangible assets and no dividend restrictions in the case of insolvency) benefited most from high CSR scores. Specifically they find that, all else equal, a one standard deviation increase in their CSR measure is associated with a 34bps reduction in credit spread. This rises to 43bps and 52bps respectively for firms that are more able to engage in either asset substitution or cash diversion when experiencing financial distress and up to 62bps for non-investment grade firms (footnote their measures). They also find that high-CSR firms were able to raise more debt at lower spreads with better initial credit ratings and longer maturities. This further supports their argument that strong CSR firms more effectively establish trust and mitigate agency frictions between contracting parties, which exerts a positive influence on credit ratings. Lastly, they do not find a similar cross-sectional relationship in equity returns, which they argue suggests that the mechanism through which ESG influences debt (agency costs of debt), is different from that through which ESG influences equity. Hoepner and Nilsson (2020) adopt a conceptual framework based almost entirely around the alignment of shareholder and bondholder interests. Unlike the previously mentioned authors, however, they suggest that, just like low levels of CSR, high levels of CSR also indicate greater misalignment between shareholders and bondholders.

3.6 Market's pricing of ESG

That bondholders perceive CSP as financially material does not in itself suggest that an ESG-based investment strategy should (or even can) deliver abnormal returns. As Goss and Roberts (2011) note, many of those who argue against ESG integration suggest it to be a futile exercise, as any corporate action (socially responsible or otherwise) will be immediately reflected in asset prices. Any observed relationship between CSR and asset returns, therefore, should disappear entirely when those returns are viewed on a risk-adjusted basis. Moreover, should ESG act primarily to reduce firm-specific risks (as well as a firm's exposure to market-wide risks) as the literature suggests, this would imply lower expected returns on assets from companies with strong ESG performance (Scharfman and Fernando). On the surface, therefore, evidence that high ESG assets generate abnormal returns (on an absolute as well as a risk-adjusted basis) appears to contradict conventional asset pricing theory (Guenster et al. 2016). This paper considers two possibilities. The first, which we will discuss in this section, is that the market does not appropriately price the economic consequences of ESG. The second, which we will discuss in the next section, is that conventional asset pricing models have been unable to account for the variation in bond returns due to CSR.

The literature focuses on two competing mechanisms through which ESG may distort prices, which are commonly referred to as the "errors-in-expectations" and the "shunned stock" effects (Derwall et. al 2011). The former suggests that the value-relevant information related to ESG is slowly incorporated into a company's stock price, enabling investors to profit from the market's undervaluation of stocks from companies with strong ESG performance. This is consistent with researchers' descriptions of ESG as a complex, multi-dimensional and partially subjective as well as evidence that its benefits are largely intangible and realized over a long time horizon. For example, Edmans (2011) attributes the positive abnormal returns he finds to stocks from companies with strong employee-relations to the market's documented misvaluation of intangible assets including patented technology (particularly software) and trademarks (Aboody and Lev, 1998, Chan et al. 2001). Furthermore, Derwall et al. (2011) argue that evidence that investors react much more strongly to negative CSR news than they do to positive news supports their view that the risks to poor CSR are more easily quantified than are the benefits to strong CSR. Crucially, the errors-in-expectations hypothesis assumes that CSR can only be a source of abnormal returns insofar as CSR practices generate profits (reduce risk) beyond what investors expect. The implication, therefore, is that any positive returns associated with ESG will disappear over time as investors learn how to appropriately value ESG information.

A number of researchers attribute diminishing returns to ESG equity portfolios to the "errors-in-expectations" effect (Borgers et al. 2013, Halbritter and Dorfleitner (2015), Badia and Cortez

(2020)). Bebchuk et al. (2013) find a significant, positive link between strong governance and both operating performance and market value. While these relationships remain stable across their sample period, they show that, consistent with greater learning and attention to good governance, investing in firms with strong governance has stopped generating abnormal returns since the early 2000s.[45] Similarly, Borgers et al. (2013) find that while trading strategies based on a stakeholder-relations index produced statistically and economically significant risk-adjusted returns between 1992 and 2004, these returns declined from 2004 to 2009. They contend that investors now fully appreciate the impact of stakeholder management on stock prices. Lastly, Badia and Cortez (2020) build responsible stock portfolios based on ESG ratings for four separate regions (including the North America., Europe, Japan and Asia Pacific) and examine their returns across three subperiods (2003–2007, 2008–2012, and 2013–2017). They find that the returns to the North American ESG portfolio started to decrease in the first subperiod, while returns to Japanese and European portfolios did not waver until the 2008 – 2012 period. The authors suggest that different regions are in various stages of development with regard to investor awareness and understanding of the valuation impact of CSR practices.

The "shunned-stock" hypothesis is built upon the literature concerning segmented capital markets and Merton's (1987) incomplete information model (Derwall et al. 2011). In this model, information asymmetry causes investors to neglect certain stocks, which because of a smaller investor base, trade a discount relative to the price implied by their fundamentals. The "shunned-stock" explanation assumes that investor preferences can cause a similar degree of market segmentation, such that investors with non-pecuniary motives create a shortage of demand for socially irresponsible assets, driving up their expected returns. In a static model in which a subset of investors have ESG preferences, Heinkel et al. (2001) show that high ESG stocks come with lower risk premia and are valued more highly. In a perfectly efficient market, investors who are not norms-constrained would arbitrage the differences in value between otherwise identical high and low ESG stocks. However, as Heinkel et al. (2001) note, market segmentation limits risk sharing, and they therefore suggest that greater idiosyncratic risk may explain the higher expected returns of shunned stocks. Lee and Faff (2009) attribute variations in idiosyncratic risk, which conventional asset pricing models do not account for, to their finding that low ESG stocks outperform high ESG stocks.[46] Similarly, Hong and Kacperczyk (2009) suggest that social norms result in heightened litigation and regulatory risk for companies operating in "sin" industries, which in conjunction with a smaller investor base, accounts for the return premium they observe in "sin" stocks. They define this as stocks of companies involved in gambling and in the production of alcohol and tobacco.

The shunned-stock hypothesis predicts that demand for ESG causes high ESG securities to outperform in the short-term, but that over long-term, neglected, socially irresponsible assets outperform both conventional and high ESG securities. Since the errors-in-expectations effect may obscure the shunned-stock effect regarding the returns to high ESG assets, researchers instead focus on the returns to socially irresponsible assets, which should increase over time along with the number of norms-constrained investors. For example, Badia and Cortez (2020) find that, in contrast to those in the U.S., Europe and Japan, low ESG stocks in the Asia Pacific region out-perform high ESG stocks. Moreover, while the positive returns to high ESG strategies in the other regions decrease over their sample, the positive returns to low ESG strategies in the AP region increase (footnote: somewhat consistent with results of Yen et al. (2019) who find that, while Japanese SRI funds outperform the market, those in the AP region do not). Additionally, the shunned-stock effect assumes that some investors

[45] This is in line with Core et al. (2006), who find that the returns to the governance measure introduced in Gompers (2003) had started to disappear only a few years after the time period his sample covered ().
[46] He also uses DJSI, which the results of Menz (2010) and Magnelli (2017) suggest may be suspect

are driven purely by non-pecuniary motives, and that such investors are sufficiently large in number so as to drive the degree of market segmentation required to impact prices (based on their model, Heinkel et al. (2001) suggest such investors should comprise at least 10% of all financial market actors). Therefore, while researchers most commonly investigate the shunned-stock/errors-in-expectations hypotheses by examining the time-series variation in the returns to ESG-sorted portfolios, evidence pertaining to investor preferences for ESG is relevant as well. For example, consistent with the assumption that financial market participants neglect stocks that violate societal norms, Hong and Kacperczyk (2009) find that institutional investors are less likely to hold sin stocks and that such stocks receive less coverage from sell-side analysts. Furthermore, Badia and Cortez (2020) note that, compared to the other regions in their sample, the AP region boasts the highest percentage of SRI assets, which they suggest indicates the presence of more norms-constrained investors and thus greater market segmentation. Some papers, however, examine the role of investor preferences more directly. For example, El Ghoul and Karoui (2017) that the relationship between fund performance and investor inflows is much weaker for SRI funds compared to conventional funds. They argue that this suggests that investors derive non-pecuniary rewards from investing in high-ESG funds. Barber et al. (2020) also find a similar result in regards to dual-objective venture capital funds, which earn lower returns than other funds, but retain their limited partners.

Lastly, Dominguez and Matallin-Saez (2016) examine the performance of the VICEX fund, a fund that invests in socially irresponsible companies that SRI screens usually exclude (e.g. alcohol, tobacco, and firearms). This fund slightly out-performs conventional funds and SRI funds over their sample period, but significantly underperforms both during the financial crisis. While they argue their results support the attractiveness of controversial stocks from the perspective of financial performance, they struggle to separate fund-specific characteristics and do little to isolate the impact of ESG from other factors. However, it does suggest an alternative explanation for the "shunned-stock" effect, such that ESG mediates a company's exposure to systemic risk. Pastor et al. (2020) develop a theoretical model with three types of investors with varying degrees of ESG preferences. Important here is that they attribute negative alphas (equity portfolios) to investor preferences and climate risk (as systematic risk factor). Bolton and Kacperczyk (2019) conclude that investors demand compensation for exposure to carbon risk in the form of higher returns on carbon-intensive firms. Krüger, Sautner, and Starks (2019) find that institutional investors consider climate risks to be important investment risks.

To our knowledge, only Pereira et al. (2019) address either effect in regards to the bond market. For a sample that covers 189 European companies between 2003 and 2016, the authors find that high ESG bonds perform no differently than do low ESG bonds. However, upon separating the performance of high and low ESG bond portfolios over time, they find that, from 2003 until 2007, investors could generate positive abnormal returns by going long high ESG bonds and short low ESG bonds. They show that the positive alphas of high ESG portfolios slightly decrease from 2003 to 2010, while the negative alphas of low ESG portfolios also decrease, and even turn positive, after 2010. They suggest that these results demonstrate the presence of both the errors-in-expectations and the shunned-stock effects in the European corporate bond market. The authors control for a host of firm and industry-level characteristics and find that their results are robust to a number of alternative specifications. However, some factors diminish the generalizability of their findings. First, their sample is relatively small given the number of different countries represented. The literature suggests significant regional variation in the performance of SRI assets (Stellner 2015), even among different Eurozone countries (Leite and Cortez 2016). Second, the combination of the financial crisis and the European sovereign debt crisis significantly impacts their results (see Leite and Cortez, 2016). Still, the trends they observe are even more distinctive than many of those in the literature concerning equities, and they certainly warrant further investigation.

The remaining papers that explicitly examine the relationship between ESG and bond returns fail to detect similar trends. Leite and Cortez (2016) and Henke (2016) examine the performance of SRI bond funds in the Eurozone over a similar period as that of Pereira et al. (2019). However, given the difficulty involved in separating the impact of various fund-level characteristics from ESG in regards to fund performance, the results of such studies are not very comparable to those of Pereira et al. (2019). Still, neither study detects similar time-series variation in the performance of SRI funds, with Henke (2016) even finding that the outperformance of SRI bond funds increased post-2007. Leite and Cortez (2016) detect significant variation in the performance of French, German and UK funds, such that UK funds significantly underperform conventional funds while German funds outperform. However, they do not find that the performance of any fund group changes significantly over the sample period. Hoepner and Nilsson (2017) employ a study structure similar to that of Pereira et al. (2019), but they do not conduct a careful investigation of the time series variation in the returns to their synthetic portfolios. They find that portfolios of bonds without ESG concerns or ESG strengths outperform the market over their full sample period (X to X). Additionally, they find that the outperformance of such portfolios was larger during the financial crisis. However, they do not examine whether this relationship was stronger in the beginning of their sample versus the end.

3.7 Demand vs. risk-based explanations for ES-spread relationship

This paper previously examined the literature whether credit spreads reflect the literature's findings concerning the economic effects of ESG. As the shunned stock/errors-in-expectations framework highlights, however, it is important to investigate whether, and to what degree, the apparent relationship is due to the pricing of fundamental, economic effects or the result of investor preferences. As Halling et al. (2021) note, compared to equity markets, bond markets have the potential to offer a much clearer setting in which to discern the effects of ESG investor preferences, and bond risk premia can be quantified via credit spreads. Halling et al. (2021) present a basic theoretical model based on Heinkel et al. (2001) to understand the role of ESG in shaping corporate credit spreads. The model assumes two types of investors, those who only buy bonds from firms with good ES-ratings and those who are neutral with respect to ES performance. As a result, in addition to changes in the probability of default and the expected losses in case of default, investor preferences shape corporate bond spreads. This simple intuition leads to a framework similar to that of the errors-in-expectations/shunned stock, which makes a number of predictions about how demand-based and risk-based mechanisms would express themselves in bond spreads.

Halling et al. (2020) find that various dimensions of ES-scores correlate strongly with both defaults and the likelihood of ratings downgrades, which further suggests a risk-based explanation for the relationship between ESG and credit risk. Much of the evidence previously covered in regards to ESG and environmental performance as well as measures of stakeholder relations implicitly support the risk-based argument, as it appears that investors discern between different categories of social performance, which are consistent with evidence concerns the relative economic benefits of those dimensions.

Barth et al. (2020) effectively isolate credit risk from other determinants of bond spreads by examining the relationship between ESG and CDS spreads. As the authors note, CDS spreads provide a comparatively precise measure of credit risk as they are less influenced by investor preferences or liquidity and have been shown to account for the majority of the firm level determinants of default risk (Tang and Yan, 2010). Barth et al. (2020) results strongly suggest that ESG practices are connected to credit risk. For a sample of 470 U.S. and Europe-based firms from 2007 to 2019, they find a significant and negative relationship between ESG and credit risk. This result is robust to controls for endogeneity and reverse causality, as well as a variety of firm and bond-level characteristics. They also find that most other determinants of CDS spreads, such as stock returns and leverage, are seemingly unrelated to ESG. They do find, however, that improvements in ESG correlate with reductions in volatility. This amplifies the risk mitigation effect of ESG by up to 25%.

3.8 Credit ratings

Credit ratings are another facet of the bond market that differs significantly from the equity market. Barth et al. (2020) are one of the few papers to examine whether credit ratings can account for their results, and find that bond markets assess the value of ESG differently than do credit ratings agencies. Credit ratings agencies are increasingly incorporating ESG considerations in their rating decisions (S&P), and Nauman (2019) expects ESG ratings and credit ratings to converge in the near future. Assuming the efficient transmission of information from credit ratings to bond prices (citation) should credit ratings account for the information reflected in ESG ratings, investors can no longer profit from such ratings. Additionally, while ratings agencies also face some social pressure to incorporate ESG issues into their assessments of credit risk (all the major agencies signed the UNPRI, also greater scrutiny following GFC), the expectation is that, like CDS spreads, credit ratings more cleanly represent credit risk than do corporate credit spreads.

Yang (2020) examines the informational value in credit ratings as well as the extent to which ESG issues predict changes in ratings. He studies the reaction of stock prices to ratings downgrades by the three major agencies (Fitch, Moody's, S&P). While S&P and Moody's announced frameworks for incorporating ESG information into their credit risk models in late 2015, Fitch did not roll out its ESG adoption until the end of 2017. The author shows that investors reacted much more strongly to downgrades from Moody's and S&P than they did regarding those from Fitch, but that this gap closed once Fitch adopted ESG. These results suggest that investors place greater confidence in credit ratings of providers who account for ESG issues. On the other hand, they are unable to predict changes in credit ratings based on ESG issues. They argue that if ratings agencies simply incorporate ESG information in an effort to portray themselves as socially progressive institutions, then the informational value of credit ratings should decline.

Lastly, the literature suggests that the relationship between ESG and credit spreads changes across the credit spectrum. If changes in credit spreads reflect the market's pricing of factors that influence default risk, then the relationship between ES-performance and spreads should be stronger for bonds that are closer to the default boundary and therefore more exposed to credit risks. However, as Halling et al. (2020) explain, demand-based effects should also be stronger for lower-rated bonds, as the risks and costs associated with ESG-arbitrage increase as credit quality deteriorates. This arbitrage is nearly riskless for highly rated bonds, since they are not subject to at least partially idiosyncratic and therefore non-hedgeable credit risks. This establishes two predictions. The first is that the negative relationship between ESG performance and credit spreads increase (decrease) as credit quality decreases (increases) (assuming of course a negative relationship between ESG-performance and credit risk, which the literature establishes). The second is that, among bonds with low-ESG scores, the positive relationship between ESG-ratings and credit spreads increase.

4. Data

4.1 Corporate bond data

In order to construct a comprehensive panel of corporate bonds, previous studies either filter massive quantities of transaction-based data (e.g. Bali et al., 2019) or rely on the constituent lists of large bond indices (e.g. Israel et al., 2018). Examples of the latter include Bloomberg Barclays indices (available on both Bloomberg and Thomson Reuters with higher tier subscriptions) and ICE BofA Fixed Income indices (which must be purchased separately). Lacking either option, we take somewhat unconventional steps to build our corporate bond dataset. Our primary sources of data include Bloomberg and Thomson Reuters. The former enables us to construct a list of prospective corporate bonds together with various bond characteristics, but its limits on data collection restrict our access to bond prices.[47] Thomson Reuters, on the other hand, does not offer comprehensive, historical lists of bonds, but provides easy access to price data.[48] In the following subsections we describe how we use these two platforms and various filtering techniques to build our corporate bond dataset.

4.1.1 Bloomberg

We use Bloomberg's SRCH function to collect identifying information on all bonds available on the platform. SRCH enables us to filter these bonds according to a variety of issue/issuer characteristics, and to specify which of these characteristics to display in exportable tables. We detail the steps we take to download this data below. The column data, however, is relatively limited, and we can only retrieve it for particular dates, the earliest of which is January 1st, 2015. Additionally, most historical market-based data are only available via the Bloomberg API, which has both a daily and monthly data limit. Since companies issue multiple bonds that differ in ways that impact returns (including, but not limited to, maturity, coupon type, series, seniority and embedded options), the volume of data required to build a corporate bond dataset dwarfs that of a comparable equity dataset. Therefore, even with a sample period of just six years, we rely on Thomson Reuters Eikon to collect price and trade-related information.

1. We filter out all bonds that are not USD denominated U.S. corporate issues. Following Bai et al. (2018), we also exclude bonds that are not listed or traded in U.S. public markets, such as those issued under the 144A or regulation S rules.
2. As has become common practice following Elton et al. (2001), we exclude puttable bonds, DRD-eligible bonds, and legally defaulted securities. We also remove structured notes, mortgage-backed, asset-backed, agency-backed, and equity-linked securities, as they have different payout characteristics to standard coupon bonds (Bektic, 2018, Bai et al. 2018). Since option features distort the calculation of returns, we simiarly exclude convertible bonds.

[47] The SRCH function, which enables us to bypass the data limits that accompany the API, does not offer historical bond prices.
[48] While the Thomson Reuters subscription available to this study does not grant access to the constituents of some of the more widely used benchmark indices from major banks, it does offer "derived historical holdings" of certain corporate bond ETFs. With these constituent lists, however, Thomson Reuters Eikon only provides RIC codes to bonds that are currently active, making these lists unsuitable for the purposes of this study.

3. Given the difficulty of tracing cash flows of floating-coupon bonds (Bai et al., 2018), we remove bonds with variable coupon rates and retain those with fixed or zero coupons.[49]
4. In an effort to ensure consistency with studies that use data from major index providers, we eliminate bonds with less than one year to maturity. That such bonds are delisted from major indices impacts their returns (Houweling and van Zundert, 2016, Israel et al., 2018, Bai et al., 2018).

For the bonds that remain after we implement the above filters, we collect data on ISINs, issue and maturity dates, first, previous and next coupon dates, the frequency of coupon payments, payment rank, ESG disclosure scores, amount outstanding and issue/issuer credit ratings from S&P, Moody's and Fitch (as well as the BBG composite, which presents an average of the ratings from major agencies). In order to circumvent the SRCH function's five-thousand row limit, we collect data in three to four separate data pulls based on different credit rating categories (e.g. AAA to A, A- to BBB and BB- and below).

4.3.2 Thomson Reuters Eikon

With the list of unique ISINs we generate from Bloomberg, we use the formula builder in Eikon's excel plug-in to download ten years of data on clean prices, accrued interest, modified duration and the market value of debt capital outstanding. To ensure that our dataset is as free of survivorship bias as possible, we use a bond's price in the month prior to its default as an estimate of the market's expected recovery rate.[50] We also examine our data for outliers, and eliminate bonds with prices below \$5 and above \$200, or amounts outstanding below 500,000 and above 10bn.[51] We then calculate each bond's monthly total return using the following formula:

$$r_{i,t} = \frac{P_{i,t} + AI_{i,t} + C_{i,t}}{P_{i,t-1} + AI_{i,t-1}} - 1$$

$$\begin{aligned}
&where, \\
&P_{i,t}\ is\ the\ clean\ price\ of\ the\ bond\ i\ at\ month\ t, \\
&AI_{i,t}\ is\ the\ accrued\ interest, \\
&and\ C_{i,t}\ is\ the\ coupon\ payment, if\ any, from\ month\ t-1\ to\ month\ t
\end{aligned}$$

The literature presents two methods of calculating corporate bond excess returns. The first involves the subtraction of a single risk-free rate from a bond's total return (commonly the one-month Treasury bill rate as in Bai et al. (2018)). The second, which we follow, subtracts the returns of

[49] According to Kumbhat et al. (2017), as of 2017 only 2% of U.S. corporate bonds outstanding had variable interest rates (compared to 85% of corporate loans). In SRCH-generated bond lists, our floating-rate coupon filter excluded at most 6% of monthly bond observations.

[50] We borrow this methodology from Houweling and van Zundert (2016). We notice that prices of defaulted (or soon to be defaulted) bonds repeat over multiple months. We therefore take the first of these as a measure of the expected recovery rate and eliminate all subsequent prices. Our method is imperfect, however, as quote-based prices for such bonds are not as accurate as transaction-based prices, and because we likely eliminate many defaulted securities in previous stages of the data gathering process (we know of no way to measure how many of these bonds the SRCH filter excludes).

[51] See Bali et al., 2019 for recommendations on how to determine outlier thresholds in corporate bond data.

duration-matched treasuries (Houweling and van Zundert, 2017, Israel et al., 2018, Bektic et al., 2018). The advantage of this method is that it effectively drives out the Term premium, which, as Houweling and van Zundert (2017) note, investors can harvest efficiently from government bonds. The authors argue that, since investors build positions in corporate bonds to earn the Default premium, researchers should consider duration-matched excess returns to evaluate unbiased bond returns. Given that the literature attributes the relationship between ESG and bond returns to credit risks (Barth et al., 2019, Halling, 2020), this return calculation is especially well suited to our study.

Unfortunately, index providers only offer duration-matched excess returns together with their constituent data, and we must therefore derive our own method of approximating them.[52] We do so with maturity-matched excess returns. We download ten years of data on treasury bonds from one month to twenty years in maturity. We then match each of our corporate bond returns to those of treasury bonds with the closest years to maturity and subtract the latter from the former. While our method is imperfect (especially with long duration bonds), we suggest it is superior to the static rate alternative.[53] After tabulating our excess bond returns we do a final check of our data, and remove bonds with monthly returns that repeat for more than two periods (Chen et al., 2007)), or that are unusually high or low (in excess of 30% in IG and 50% in HY).

4.3.3 Representative bond filter

Lastly, we apply a filter from Haesen et al. (2013) and Israel et al. (2018) to our preliminary dataset to select a representative bond from each issuer each month. In addition to creating a sample of liquid and cross-sectionally comparable bonds, this filter helps us to reduce the volume of data we require and lowers the likelihood that we trigger additional limits on data retrieval.[54]

1. We remove all non-senior debt, which have different payout characteristics compared to standard senior coupon bonds.[55]
2. We exclude bonds with credit ratings that diverge from that of their issuer.[56]
3. We retain only bonds with a time to maturity between five and fifteen years. In cases where an issuer has no such bonds, we keep all bonds from that issuer.
4. We exclude bonds that are older than two years (unless an issuer has no such bonds). This filter is based on the observation that a bond's trading activity falls substantially two years after it is issued.[57]

[52] Duration or maturity matched spreads and returns are also available with higher tier subscriptions to Thomson Reuters Eikon.
[53] Maturity and duration diverge most in long duration bonds. As a result, a positively sloped yield curve causes us to slightly underestimate the returns of a long duration bond, while a negatively sloped curve causes us to overestimate these returns.
[54] With respect to both Bloomberg and Thomson Reuters terminals, these limits are nebulous, and to our knowledge, there is no way to determine how close a user is to the data limit.
[55] Although we would have preferred to implement this filter during the Bloomberg stage in order to reduce the volume of data we require, we discovered that, with respect to payment rank, the SRCH function is often inaccurate.
[56] This step deviates slightly from that of Israel et al. (2018), who only keep bonds belonging to the rating category that contains the largest fraction of a company's total debt outstanding. The authors note, however, that in a vast majority of cases, these bonds have the same rating as their issuer.
[57] A bond's age percent, measured as a bond's time since issuance divided by its original maturity, is often used as a measure of bond illiquidity (Israel et al., 2018).

5. After applying the above filters, an issuer's representative bond is that which has the largest amount outstanding.

Our final corporate bond dataset runs from January, 1st , 2015 to March 1st, 2021, and includes 85000 bond month observations, which translates to 1896 unique corporate bonds from 1275 issuers (of which ~66% are IG and ~34% are HY).[58] The average month contains 750 bonds representing $549bn of total notional outstanding. In table 1 we display some descriptive statistics. To construct the numerical credit ratings that we use in later regressions, we map categorical ratings from AAA to D to scores from 1 to 21. Given how sparse our data is for the very highest and lowest rated bonds, we combine AAA and AA+ into one numerical score (1), as well as ratings below C (21). Otherwise, we match numbers directly to letter grades in ascending order, such that scores from 1 (AAA/AA+) to 9 (BBB-) correspond to IG, and scores of 10 (BB+) and above correspond to HY. The average issue in our sample has a rating of BBB-, a duration of 6.11 and $541 million of notional outstanding. The average monthly excess return of 0.32% and standard deviation of 3.29% is comparable to that of Israel et al. (2018) (Table 2, pg. 23).[59]

Table 1 Overview of Corporate Bond Data Pre-ESG:

Before ESG Filter	**mean**	**std**	**5%**	**25%**	**50%**	**75%**	**95%**
Credit Rating	9.34	2.79	5.00	8.00	9.00	10.00	15.00
Amt Out	541.7	545.4	44.4	218.2	400.1	674.2	1522.1
Years to Maturity	8.52	6.96	1.67	4.29	7.04	11.13	20.39
Duration	6.11	3.16	1.56	3.78	5.82	7.97	12.01
Excess Returns	0.32%	3.29%	-2.94%	-0.47%	0.21%	1.13%	3.90%
Momentum	1.92%	7.31%	-5.51%	-0.65%	1.49%	4.16%	11.02%
VaR 5%	-3.45%	2.99%	-0.89%	-1.81%	-2.76%	-4.03%	-8.50%

	Unique Bonds		Unique Issuers	
Before ESG Filter				
IG	1251	66%	804	63%
HY	645	34%	471	37%
Total	1896	100%	1275	100%

It is very important to scrutinize the quality of our dataset as our process of data collection differs from some of the highest caliber publications on corporate bonds (Israel et al., 2018, Bai et al., 2018, Bali et al., 2019). While we cannot be certain, there are two rough methods that we use to assess the quality of our data. First, we compare the table below to that of Bai et al. (2018) (Panel A, pg. 12), who use high quality, transaction-based data on U.S. corporate bonds from 2002 to 2015. Based on this high level analysis, we detect only minor differences, which we expect given our representative bond filter and different sample period.[60] The other comparison we make concerns the

[58] This ratio between IG and HY is approximately in line with those of Houweling et al., 2017, Israel et al., 2018 and Bai et al., 2018.
[59] After we account for the average market return during our sample, which is nearly double that of the authors (their sample runs from 1997 – 2015).
[60] In terms of minor differences, the distribution of average years to maturity and duration in our data is narrower (as we would expect based on our filter), and our average returns and amounts outstanding are slightly higher (which we also expect given the rise in aggregate corporate debt and higher average returns during the 2015 – 2021 period compared to those of 2002 – 2015). Bai et al. (2018) simply subtract the one-month treasury bill rate from a bond's monthly total return to calculate its excess return.

performance of our factor-mimicking portfolios relative to those of Houweling and van Zundert (2017). As we discuss in more detail in the following section, our results match theirs very closely.

4.3.4 Limitations

While invaluable in terms of the access they grant to a broad range of bond and firm-level data, our reliance on student licenses to Bloomberg and Thomson Reuters Eikon has a number of drawbacks that may limit the precision and generalizability of our findings. These result primarily from our use of quote-based prices and our relatively narrow sample period. Regarding the former, Bessembinder et al. (2006, 2009) emphasize the importance of using transaction-based data from the Trade Reporting and Compliance Engine (TRACE) over dealer-quoted data that is available from major vendors such as Bloomberg or Thomson Reuters. The authors show that the explanatory power of test statistics designed to detect abnormal bond returns in corporate event studies increases when using daily and monthly returns calculated from TRACE data. Other studies that investigate common factors in corporate bond returns follow Bessembinder et al. (2006, 2009) in using transaction-based data (Jostova et al. (2013), Chordia et al., 2017; and Choi and Kim, 2017, Bai et al., 2018). Unfortunately, this data is not easily accessible, and as a result, we rely on quoted prices from Thomson Reuters. This is not unprecedented, however, and to our knowledge all of the studies that use synthetic portfolios to examine the relationship between sustainability and corporate bond returns do so with quote-based prices (Zerbib 2018, Pereira et al. 2019, Hoepner and Nilsson, 2020). Additionally, in an effort to expand their analysis to a period prior to the introduction of TRACE data, Bai et al. (2018) successfully replicate their main results with quoted prices.

As in Israel et al. (2018), a consequence of our bond filter is that we do not examine a liquidity premium. Unlike the authors, however, it was not our intention to omit a separate liquidity factor. Bao et al. (2011) and Bai et. al. (2018) highlight the inability of researchers to effectively isolate a liquidity premium without access to daily, transaction-based prices. Slimane et al. (2019) note that the difficulty involved in the acquisition of such data means that investors less frequently integrate liquidity into their bond risk models. While many studies (predominantly those in the first decade of the new millennium) emphasize liquidity's importance in explaining the cross-sectional differences in credit spreads (Longstaff et al., 2005, Lin et al., 2011), Slimane et al. (2019) find liquidity factors to be among the weakest in explaining Euro-area corporate bond returns from 2009 to 2018. This is in line with Chen et al. (2007) who find that differences in maturity, bond size and credit quality can explain most of the cross-sectional differences in trading costs. Furthermore, the literature identifies a range of effective liquidity proxies, such and bond size and bond age, which we capture in our other factors (Sarig and Warga, 1989, Crabbe and Turner, 1995 and Houweling and van Zundert, 2017). Thus, while we do not include a liquidity factor in our performance evaluation models, we suggest that our Size factor, controls for various bond characteristics and bond filter help us to account for the effects of a potential liquidity premium in the returns to our bond portfolios.[61]

To some extent, our short sample period undermines our ability to evaluate the time dependency of returns to ESG-sorted portfolios as well as the validity of our results more broadly. Unfortunately, it is not possible with Bloomberg's SRCH function to prepare a list of representative bonds prior to January 1st, 2015. We explored alternative methods of extending our dataset, such as gathering additional data on bonds in our master list that were issued prior to the cut-off date. Although this

[61] Our representative bond filter provides us with what is essentially a liquid subsample of bonds. Houweling and van Zundert, 2017 (pg. 27) employ a similar filter to examine whether the performance of their factor portfolios remains robust in a liquid subsample.

provides us with an additional five years of data (which becomes unbearably sparse if we go back more than five years), the extended panel of bonds is not comparable to our original panel, which complicates any inferences we can make from it. We therefore decide to follow Nagy et al. (2016), who intentionally limit their sample to eight years in order to have a more homogenous dataset in terms of asset coverage and methodological consistency. Moreover, while it has no bearing on the validity of our results, we note that other recent studies examine the relationship between ESG and corporate bonds with comparably short samples (Magenelli and Izzo 2017, Cubas-Díaz and Ángel Martínez Sedano 2018).

4.2 ESG data

Our main ESG dataset comes from Thomson Reuters, whose current ESG scores replace the popular ASSET4 ratings.[62] Many previous studies have relied on ASSET4 ESG scores, and Stellner et al. (2015), Eding and Scholtens (2017), and Pereira (2019) highlight the advantages these scores offer in terms of both company coverage as well as consistency and reliability. To justify their use of the new Thomson Reuters ESG scores (henceforth TR scores), Dorfleitner et al. (2020) cite Thomson Reuters' transparent scoring methodology and its status as the world's largest ESG ratings database. TR scores are also widely available and frequently used in practice, which enables us to both push back against researchers' use of proprietary (or otherwise restricted) ESG ratings (Bennani et al. (2019)), and to better capture potential demand-based effects. Iaonnou and Serafeim (2012) estimate that investors who use ASSET4 ratings to integrate ESG data into the portfolio construction process represent upwards of 2.5 trillion euros under management. However, as Pereira (2019) notes, one downside of TR scores is that Thomson Reuters/Refinitiv revises them retroactively. While these revisions seem infrequent, there is no way for us to retrieve a company's TR score as it appeared at a specific point in time. It is therefore unclear to what extent Thomson Reuters amended the previous ASSET4 ratings after it implemented its enhanced scoring methodology. After a close comparison of the ASSET4 framework to the current system, we believe that, while Thomson Reuters almost certainly revised earlier scores (partly due to the dissolution of ASSET4's economic pillar, which comprised a quarter of its equally-weighted ESG score), the criteria are so similar that whatever trends the previous scores exhibited are reflected in the current scores as well.

To construct its core ESG score, Thomson Reuters assesses a company's environmental, social and corporate governance performance separately. These three pillars of a company's ESG performance encompass ten themes based on over 400 data points and 70 performance indicators (for a more detailed explanation see Thomson Reuters, 2018). Thomson Reuters assigns category scores based on a firm's performance across each theme. The environmental pillar represents themes of resource use, emissions and product innovation, while the social pillar covers human rights, community, workforce and product responsibility, and the governance pillar reflects CSR strategy, management and shareholders. Scores are based on publicly reported data, and reflect both a company's performance in each category as well as its transparency.[63] Thomson Reuters also incorporates adjustments at the category level based on materiality, industry relevance and data availability. Scores are updated on a biweekly basis, and although Thomson Reuters issues new scores annually (as is the

[62] Thomson Reuters note that its current ESG scores are an "enhancement and replacement to the existing ASSET4 ratings" (Thomson Reuters, 2018, pg. 3). Although we keep referring to Thomson Reuters as the provider of these scores, Refinitiv has officially published them since 2018.

[63] Thomson Reuters reduced the importance of ESG disclosure relative to ESG performance in its new scoring framework (Thomson Reuters, 2018, pg. 3).

case for other major index providers), it promptly adjusts a company's score if new information alters its assessment. In our data, we find that 7% of covered companies experience intra-annual score changes. Lastly, Thomson Reuters combines a company's category scores and assigns a percentile rank for each pillar, which corresponds to a letter grade between D- and A+, such that a score between 66.6 and 75 equates to a B+ (see Thomson Reuters, 2018, pg. 7). A company's overall ESG score similarly reflects a percentile rank, which Thomson Reuters calculates after combining each company's pillar scores. Importantly, Thomson Reuters calculates these percentile ranks within each industry group separately, a process Dorfleitner et al. (2020) find is sufficient to control for the variable effects of industry membership on the ESG-CFP relationship (see Edmans, 2011 or Gregory et al., 2016 for an overview of how industry membership can influence ESG scores).[64] We also download data on Thomson Reuters' ESG controversies score and the Bloomberg ESG disclosure score, both of which we discuss in Section 6.3c.

We download ten years of ESG scores on all U.S. companies (3055 in total) in the TR ESG database (which includes Environmental, Social and Governance pillar scores, Combined scores and ESG scores with a controversies overlay). We then use the formula builder to retrieve Thomson Reuter's "ultimate parent" identification code for each unique ISIN in our corporate bond sample. Lastly, we use this code to match each bond with ESG data from its respective issuer. We are able to match approximately 72% of our IG bonds and 68% of our HY bonds. In table 2 below, we summarize the characteristics of our dataset after matching our bonds with TR scores.

Table 2 Overview of Corporate Bond Data with ESG:

After ESG Filter	**mean**	**std**	**5%**	**25%**	**50%**	**75%**	**95%**
Credit Rating	9.13	2.48	5.00	8.00	9.00	10.00	14.00
Amt Out	541.6	542.5	46.5	228.8	396.3	672.9	1529.5
Years to Maturity	8.61	7.31	1.67	4.38	7.13	11.21	20.39
Duration	6.12	3.12	1.55	3.82	5.88	8.01	11.81
Excess Returns	0.31%	3.07%	-2.81%	-0.47%	0.20%	1.11%	3.75%
6 m Excess Returns	1.83%	6.94%	-5.21%	-0.63%	1.42%	4.01%	10.29%
VaR 5%	-3.28%	2.87%	-7.17%	-3.87%	-2.66%	-1.80%	-0.89%
ESG	58.1	18.0	25.0	44.5	61.0	71.9	83.5
ESG Combined	52.8	17.0	24.2	40.5	52.7	65.6	80.1
Environment	50.2	28.3	0.0	27.4	55.2	74.8	87.4
Social	59.7	20.3	24.2	44.3	62.5	75.8	89.3
Governance	61.9	19.9	23.8	49.2	64.7	77.4	89.2
ESG MOM	8.51%	212.95%	-11.34%	-1.11%	1.66%	9.66%	31.89%

	Unique Bonds		Unique Issuers	
After ESG Filter				
IG	971	70%	635	68%
HY	424	30%	294	32%
Total	1395	100%	929	100%

[64] Halling et al. (2020) find that, among U.S. bonds, industry membership has no effect on the relationship between ESG and credit spreads.

5. Methodology

5.1 Portfolio Construction

5.1.1 Separation of investment grade and high-yield bonds

The notion that the behaviors of investment grade (IG) and high-yield (HY) bonds differ systematically dates back to Merton (1974), whose structural model of credit risk predicts that bonds with a lower distance to default become more equity-like. Derwall and Koedijk (2009) similarly find that HY bonds exhibit a degree of firm-specific risks more similar to equities than to IG bonds, which they suggest provides investment managers with the opportunity to better exploit ESG information in HY bond portfolios. Increasingly, however, researchers find that the relationships between credit ratings, risk, returns and other bond characteristics are not linear, but rather reflect two distinct markets divided between the IG and HY boundary. Chen et al. (2007) find that IG and HY bonds exhibit different exposures to liquidity risk and transaction costs, while Ambastha et al. (2010) show a sharp decline in interest-rate sensitivity from BBB to BB rated bonds. Chen et al. (2014) provide an overview of the literature on the relationship between labels and security prices, which finds that the former can influence the latter through its impact on investors' preferences and holdings. They offer strong, empirical evidence that such a label effect exists with respect to credit ratings, such that ratings divide the bond market between IG and HY bonds.

We separate IG and HY bonds in our construction of both factor and ESG portfolios. Most authors only split their sample this way to control for differences in credit ratings (i.e. Bai et al., 2018), but others apply this separation throughout their analyses (Houweling and van Zundert, 2017, Bektic et al., 2018). We follow the latter, both due to evidence of market segmentation, as well as the practice of asset owners, index providers, investment managers and regulators who treat IG and HY as two separate asset classes (Houweling and van Zundert (2017)). As we discuss throughout Section 6, our results further emphasize the importance of developing separate IG and HY benchmark and factor portfolios. To our knowledge, we are the first paper to apply this distinction in an examination of ESG-sorted corporate bond portfolios.[65]

5.1.2 Cut-off rates and weighting schemes

Many researchers highlight the pitfalls of cap-weighted portfolios in general (Fama and French, 2008) and in the corporate bond market in particular (Siegel, 2003, Baker and Wurgler, 2012, Choi and Kim, 2016, Bektic et al., 2018). To ensure that single, large issuers do not dominate our portfolios, we construct equally-weighted benchmark, factor and ESG portfolios (Houweling and van Zundert, 2017, Bektic et al., 2018). That each bond receives an equal weight also means that we explore a potential size premium separately. Fama and French (2008) also note potential problems associated with equally-weighted portfolios in overrepresenting the smallest issuers (or micro caps). This may be especially problematic for bonds, as the smallest issues carry, on average, larger transaction

[65] Although Hoepner and Nilsson (2020) control for differences in credit ratings, they implement these controls across the ratings spectrum, which obfuscates any potential effects attributable to the IG/HY divide. Most important, however, is that we develop two sets of multifactor models where most papers rely on just one (Bai et al., 2018, Hoepner and Nilsson, 2020).

costs, and may be untradeable for most investors. We implement a variety of controls to minimize the influence of the smallest bonds on our results, such as excluding bonds with less than 25m USD notional outstanding, and splitting our sample between small and large bonds. We are less concerned about our weighting scheme in regards to our factor-mimicking portfolios as we borrow our methodology from the literature. Additionally, Houweling and van Zundert (2017), Israel et al. (2018) and Bai et al., (2019) find their results to be consistent across equal- and value-weighted portfolios. With respect to ESG portfolios, however, we examine the role of bond size more carefully. Although many studies regarding the relationship between ESG and financial returns arrive at similar results with equal and value-weighted portfolios (Kempf and Ostoff, 2007, Dorfleitner and Hallbritter, 2015, Hoepner and Nilsson, 2020), Statman and Glushkov (2009) and Dorfleitner et al. (2020) notice small differences, which Dorfleitner et al. (2020) attributes to the influence of company size on the market's perception of the marginal benefits of ESG.

Additional considerations include portfolio cut-off rates and transaction costs. It is a common practice in the literature to investigate the existence of factor premia via either decile or quintile analysis (see Benartzi et al., 1997, Frazzini and Pedersen, 2014 or Bektic et al., 2018). We rank and group bonds into decile portfolios according to their factor scores in order to be consistent with the papers from which we borrow factor definitions (Houweling and van Zundert, 2017, Bai et al., 2018). Regarding ESG-sorted portfolios, on the other hand, we use quintile analysis, as the inclusion of ESG ratings reduces the number of bonds available each month. The most notable consequence of our smaller sample size is that, with a 10% cut-off rate, we would be unable to effectively construct double- and triple-sorted portfolios, which we use to control for various bond characteristics (Section 6.3c). Where possible, however, we test the robustness of our results to alternative cut-off rates of 10% and 50%.

We follow Houweling and van Zundert (2017) and Israel et al. (2018) in combining portfolio turnover with transaction cost estimates from Chen et al. (2007), which the authors present based on a bond's size, credit rating and maturity. While we do not explicitly account for transaction costs in our long-short factor-mimicking portfolios (which therefore fall under the oft-maligned category of "academic factor", Blitz et al. (2019)), we do so when we analyze long-only alternatives (see Section 6.1). Similarly, we only incorporate transaction costs in our final analysis of the top performing ESG-integrated portfolios (Section 6.5). As has become standard in the extant literature, we rebalance our portfolios on a monthly basis (Houweling and van Zundert (2017), Bai et al. (2018) and Bektic (2018)). Lastly, we lag firm characteristics one month in order to reduce endogeneity problems and to account for publication lag and the time required for market participants to incorporate firm data into bond prices (Oikonomou et al., 2014, Bektic, 2018).

5.2 Factor portfolio formation

We first calculate factor scores for each bond in our IG and HY universes. For Size, Downside-risk, Credit Quality and Momentum this is straightforward, as a bond's factor score simply reflects its amount outstanding, 5% value-at-risk, credit rating and past returns respectively. Our primary Momentum score is that of Jostova et al. (2013) and Houweling and van Zundert (2017), who define Momentum as a bond's six-month excess return with a one-month implementation delay.[66] With the exception of Low-risk, we then sort bonds according to their factor scores and group them into decile portfolios. Regarding our Size and Downside-risk scores, we rank bonds from smallest to largest, while

[66] We form Momentum on three and twelve month returns as well. This provides us with an easy to way to check if our results our robust to alternative factor definitions, something we cannot do for Value, Size and Low-risk.

for our Credit Quality, Momentum and Value scores we rank bonds from largest to smallest. The top Size decile portfolio therefore contains the 10% smallest bonds in terms of total notional outstanding, while the top Momentum decile contains bonds with the 10% highest six-month returns. Our factor-mimicking portfolios are long-short portfolios, which fund a long position in the top decile with a short position in the bottom decile.

Our factor scores for Value and Low-risk require additional computation. Regarding the former, we run a regression of credit spreads against credit rating dummies, time-to-maturity and the three-month change in credit spreads.

$$S_i = \alpha + \sum_{r=1}^{9} \beta_r I_{i,r} + \gamma M_i + \delta \Delta S_i + \varepsilon_i$$

Where S_i is the credit spread of bond i, $I_{i,r}$ is a dummy variable that takes the value 1 if bond i has rating r (0 otherwise), M_i is years to maturity and ΔS_i is change in the credit spread. Our Value factor score is equal to the difference between bond i's actual spread and the spread predicted by the model above. The higher the residual, the "cheaper" the bond is relative to its credit rating, maturity and spread change. As we explain in greater detail in 6.1, we find that the average excess and risk-adjusted returns of Value decile portfolios increase steadily from the lowest (negative) residuals to the highest, with the top three decile portfolios in both IG and HY producing Sharpe ratios twice that of the market portfolio.

We define our Low-risk factor based on both a bond's duration and credit rating (we borrow this definition from Houweling and van Zundert (2017), who in turn borrow from Illmanen (2011)). In the equities literature, most researchers define Low-risk in terms of either a security's market beta (Frazzini and Pedersen, 2014) or other measures its volatility (Maguire et al., 2017). However, since a bond's volatility approaches zero as it nears maturity, these measures are inadequate in corporate bonds (Israel et al. (2018)). To construct our long portfolio, we limit our IG (HY) sample to bonds with credit ratings between AAA and A- (BB+ to B-). Of the remaining bonds we select those with the lowest durations, such that our long portfolio contains 10% of all bonds each month. Inversely, our IG (HY) short portfolio includes the longest duration bonds rated below AA- (BB-). While this procedure ensures that our Low-risk portfolio is comparable to our other factor-mimicking portfolios, it prohibits us from examining cross-sectional trends via decile analysis, and also from implementing controls with regards to credit ratings and maturity. We find that, while our long portfolios exhibit significantly lower average excess returns compared to our short portfolios (1.97% vs. 5.43% in IG and 2.22% vs. 7.63% in HY), they produce higher Sharpe ratios (0.99 vs. 0.58 and 0.61 vs. 0.51). These results support the presence of a Low-risk anomaly in corporate bonds similar to that which Frazzini and Pedersen (2014) describe with respect to equities.

To be as consistent with Bai et al. (2018) as possible, we use the authors' double sorting procedure to construct Downside-risk and Credit Quality factor portfolios. For Downside-risk, we first group bonds into quintiles according to credit ratings, and select from each quintile the top 10% of bonds in terms of 5% VaR. To form our Credit Quality portfolio we sort bonds into quintiles based on 5% VaR and then into deciles with respect to credit ratings. Bai et al. (2018) include additional sorts for each factor in their four-factor model, and, divorced from this setting, our double sorting procedure loses much of its relevance. Although we do not take similar steps to minimize the potential influence of each factor on one another, we do examine the relationships between factors and use alternative techniques to minimize redundancies and unwanted factor exposures (Section 6.2a).

5.3 Comparing Factor Portfolio Returns

In line with the existing literature, we measure the risk and performance of each of our factor portfolios in three ways (see, in particular, Bektic et al., 2017, others include Fama and French, 1993, Gebhardt et al., 2005, Correia et al., 2012, Jostova et al., 2013, Houweling and van Zundert, 2017, Bektic, 2018).

First, we use the Sharpe ratio to measure the risk-adjusted returns of each factor and ESG-sorted portfolio *i:*

$$S_i = \frac{E[R_i - R_f]}{\sigma_i}$$

where S_i is portfolio *i*'s annual average excess return divided by the standard deviation of its excess returns.

Next, we regress the returns of portfolio *i* against those of the market portfolio (or Default premium). The intercept from this regression is the equivalent of the CAPM-alpha for the corporate bond market, and gives us an indication of how well our factor and ESG-sorted portfolios perform relative to the market portfolio.

$$R_{p,t} - R_{f,t} = \alpha_i + \beta_{1,i} DEF_t + \varepsilon_{i,t} \tag{1}$$

Many studies define *DEF* as the returns of an aggregate corporate bond index (Bai et al., 2018, Hoepner and Nilsson, 2020). We follow Houweling and van Zundert (2017) and Bektic (2018) in replacing this traditional definition with the returns of an equally-weighted portfolio of all bonds in our sample (as with our factor portfolios, we create market portfolios for IG and HY separately).

Finally, we implement our multifactor models in order to correct for multiple sources of systematic risk. The primary model we use to analyze our factor-mimicking portfolios contains the original Fama and French (1993) five-factor model (with our modified *DEF* factor) together with the equity momentum factor *UMD* (Carhart, 1997).

$$R_{p,t} - R_{f,t} = \alpha_i + \beta_{1,i} RMRF_t + \beta_{2,i} SMB_t + \beta_{3,i} HML_t + \beta_{4,i} UMD_t + \beta_{5,i} TERM_t + \beta_{6,i} DEF_t + \varepsilon_{i,t} \tag{2}$$

Like Houweling and van Zundert (2017), we define the *TERM* factor as the excess returns of a medium-term treasury bond index, which has an average maturity closer to that of our representative bond sample (compared to that of a long-term index, which other researchers seem to favor). A combination of equity factors and bond market indexes is common in the literature (Houweling and van Zundert, 2017, Israel et al., 2018, Bektic 2017, 2018, Bai et al., 2018). Some also borrow from Chen et al. (1986) and include macroeconomic variables such as the change in industrial production and inflation (Correira et al., 2012). Though we do not do so, Henke (2016) and Hoepner et al. (2018) show that *TERM* and *DEF* subsume the variation attributable to these factors in the time series. Others add equity factors like *QMJ* (Israel et al., 2018), but the marginal contribution of factors constructed from the cross-section of equity returns is minimal (Asness et al., 2014).

We apply Breusch and Pagan (1979), Breusch (1978) and Durbin and Watson (1971) tests to all portfolios. The former indicates that our residuals are subject to heteroskedasticity while the latter two show modest autocorrelation across most of our models. We therefore follow the approach of Newey and West (1987) and calculate heteroskedastic and autocorrelation robust standard errors.

6. Results and analysis

6.1 Factor portfolio analysis

We start our empirical analysis with an examination of our long-short factor portfolios. In Table 3 we report each portfolio's alpha (and adjusted t-statistic), beta (with respect to the market portfolio) and adjusted R-squared with regards to both our CAPM-equivalent model, and our Fama and French (1993) inspired multifactor model (Model 1). Unless we mention otherwise, the statistics we describe below are in reference to our results from Model 1.

Our Value portfolios produce the largest and most statistically significant alphas (3.34% (3.17) in IG and 7.67% (6.00) in HY). Across Value-sorted decile portfolios in HY, Sharpe ratios increase steadily from the bottom portfolio to the top. Although present, this trend is less defined in IG, with the top four decile portfolios showing comparable risk-adjusted performance (Sharpe ratios of ~1.2). At cut-off rates of 20% and 50% the negative returns of the bottom portfolios vanish, but the Value effect remains highly potent nonetheless. Our results are consistent with those of Houweling and van Zundert (2017), Israel et al. (2018) and Slimane et al. (2018), which find that Value outperforms other factor portfolios in corporate bonds. Israel et al. (2018) go one step further and show that a slightly modified Value portfolio displays impressive explanatory power in the cross-section of average monthly returns.

The behavior of our Momentum portfolios also aligns with those from the existing literature. Like Houweling and van Zundert (2017), we detect a momentum effect in HY and a reversal effect in IG. Among our HY factor portfolios, Momentum produces the second largest alpha (7.45% (2.01)), which comfortably exceeds that of the authors' Momentum portfolio (4.84% (2.24)).[67] On the other hand, our IG Momentum portfolio underperforms its counterpart in Houweling and van Zundert (2017) (-2.34% (-2.40) vs. -1.35% (-1.34)), and reflects a more pronounced reversal effect.[68] Popsil and Zhang (2010) and Jostova et al. (2013) discover a similar dichotomy between IG and HY portfolios, and, across integrated corporate bond portfolios, Bai et al. (2018) and Bali et al. (2019) identify a sharp reversal effect at short time horizons (one to three months).[69] We are unaware, however, of a study that finds evidence of a robust reversal effect over a twelve-month period as we do. Importantly, this effect is not the result of a rogue decile portfolio, but is rather reflected in the steady increase in excess returns and Sharpe ratios from the top IG Momentum decile (average annual excess returns of 1.82% and Sharpe ratio of 0.27) to the bottom decile (returns of 6.69% and Sharpe ratio of 1.04). In order to test whether the reversal effect is a product of our factor definition, we form Momentum portfolios with alternative lookback periods of one, three and six months. Each portfolio exhibits reversal, but we find the effect to be considerably stronger at shorter intervals, such that the portfolio sorted on monthly returns produces an alpha nearly twice that of the portfolio sorted on yearly returns (-4.48% (-2.38) vs. -2.34% (-1.79)).

[67] Which we argue is comparable as we use the same multifactor model to evaluate portfolio performance.
[68] In other words, our Momentum portfolio fares more poorly against our benchmarks compared to the performance of the authors' portfolio relative to its benchmarks.
[69] The results of Popsil and Zhang (2010) and Jostova et al. (2013) suggest the absence of a momentum effect in IG more so than the presence of a reversal effect.

Table 3: Performance of long-short factor portfolios

This table shows the performance statistics of Size, Value, 1-month Momentum, 6-month Momentum, Low-Risk, Downside, and Credit Rating factors for U.S. Investment Grade and U.S. High Yield corporate bonds over the period January 2015 to March 2021. See Section 5.2 for details regarding the construction of each factor portfolio. Panel A shows the CAPM-alpha and -beta with respect to the corporate bond market (DEF). Panel B shows alpha statistics with respect to a multi-factor model (RMRF, SMB, HML, MOM, TERM and DEF). Panel C shows correlations between the CAPM-alphas of the factors. We annualize all alphas and report the models' adjusted R^2. *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence levels, respectively, of two-sided tests on whether the alphas are different from 0 (*t*-tests with Newey-West standard errors). We measure corporate bond returns as excess returns versus maturity-matched Treasuries.

	Investment Grade							*High Yield*						
	Size	Value	Mom (1 yr)	Mom (6 mo)	Low Risk	Downside Risk	Credit	Size	Value	Mom (1 yr)	Mom (6 mo)	Low Risk	Downside Risk	Credit
Panel A: CAPM Results														
alpha	0.09%	2.07%**	-2.95%***	-2.97%*	1.04%*	-2.52%**	-0.21%	4.35%*	9.78%***	6.95%*	6.29%	0.42%	-4.71%	-6.84%
t-value	0.200	2.471	-2.981	-1.768	1.835	-2.044	-0.184	1.945	5.686	1.943	1.169	0.319	-1.466	-1.452
beta	0.134	-0.990	-0.232	0.182	-0.776	1.189	0.210	-0.443	-1.740	-1.087	-1.771	-0.577	1.136	0.988
adj R^2	0.048	0.435	0.017	0.012	0.857	0.808	0.224	0.347	0.864	0.311	0.561	0.596	0.584	0.300
Panel B: Model 1														
alpha	0.44%	3.34%***	-2.34%**	-3.13%*	1.07%**	-2.49%**	0.06%	4.94%**	7.67%***	7.45%**	8.82%	0.42%	-4.28%	-7.52%*
t-value	1.020	3.172	-2.402	-1.891	1.943	-2.057	0.077	2.214	5.997	2.091	1.543	0.458	-1.379	-1.983
adj R^2	0.097	0.499	-0.019	0.029	0.857	0.797	0.315	0.312	0.869	0.266	0.547	0.613	0.570	0.303
Panel C: Correlations														
Size		-0.23	0.36	0.43	-0.23	0.31	-0.14		0.45	0.32	0.54	0.47	-0.40	-0.35
Value			0.12	0.22	-0.67	0.75	0.65			-0.03	-0.38	-0.56	0.65	0.53
Mom (1 mo)				0.68	0.22	-0.09	-0.21				0.70	0.33	-0.33	-0.17
Mom (6 mo)					-0.17	0.19	-0.18					0.57	-0.61	-0.40
Low Risk						-0.95	-0.55						-0.82	-0.87
Downside							0.47							0.73

Our Size portfolios also indicate a rift between IG and HY. In IG we observe relatively little variation in the excess and risk-adjusted returns of decile portfolios sorted on average bond size. The top Size decile exhibits higher returns (3.97% vs. 3.32%) and a slightly higher Sharpe ratio (0.84 and 0.81) compared to the bottom decile, but these differences are not statistically significant. As a result, our IG Size portfolio fails to produce a robust intercept term. This is not the case in HY, where the top Size decile markedly outperforms the bottom decile (Sharpe ratio of 0.80 and 0.59 respectively), and our HY Size factor generates a strong alpha of 4.94% (2.21). These results broadly track those of Houweling and van Zundert (2017) and Slimane et al. (2018), both of which attribute Size's stronger performance in HY to the greater relevance of liquidity and default risk premia.[70] That we do not observe a consistent pattern in the performances of our HY Size deciles, however, undermines this explanation. Most notably, the best performing decile (decile 4, which exhibits average excess returns of 9.67% and a Sharpe ratio of 0.96) sits adjacent to the weakest decile (decile 3, excess returns of 4.69% and a Sharpe ratio of 0.38). Given the seemingly rational behavior of our other factor portfolios, we suggest this is unlikely to be simply an artefact of our limited HY dataset. We propose instead that it indicates the importance of variables other than liquidity and/or default risks in determining the relationship between average bond size and returns in HY (or, alternatively, that bond size has a limited influence on a portfolio's exposure to these risks).

The performances of Low-risk, Downside risk (DR) and Credit Quality (CQ) are less noteworthy. While our top Low-risk portfolios generate much lower average excess returns compared to our bottom portfolios (which makes sense given that they are comprised of bonds with shortest maturities and highest ratings), they boast markedly higher Sharpe ratios. Ultimately, however, it is only in IG that Low-risk produces a statistically significant alpha (1.07% (1.94)). These rather muted results align more closely with those of Slimane et al. (2018) than those of Houweling and van Zundert (2017), which could indicate that Low-risk has been less relevant in recent years.[71] Downside risk is interesting insofar as we generate the opposite result to that of Bai et al. (2018). Where the authors find a positive relationship between a portfolio's risk-adjusted returns and its VaR, we find that, in both IG and HY, Sharpe ratios decrease monotonically from top DR deciles to bottom DR deciles. Downside risk and Low-risk are both essentially beta proxies and are highly correlated with one another (-0.95 and -0.82 in IG and HY respectively). It makes sense, therefore, that our detection of a Low-risk effect precludes the presence of a Downside risk effect. Lastly, our HY Credit portfolio exhibits a strong, negative alpha (-7.6% (-1.98)), while our IG Credit portfolio fails to produce one that is statistically significant. We are not aware of prior research that can help us to reconcile these results, and, given the high variability in returns across our Credit-sorted quintile portfolios, we are inclined to blame our short sample as well as the untested nature of the CQ factor (which only appears in Bai et al. (2018)).

We find it remarkable that, although our respective sample periods do not overlap, our results are so similar to those of Houweling and van Zundert (2017). Our factor portfolios not only generate comparable returns, but our CAPM and multivariate models also explain these returns to a similar degree. Even our HY Low-risk portfolio, the alpha of which diverges most from that of the authors, shows a similar beta coefficient with respect to the market portfolio (-0.58 vs. -0.76) and an equivalent adjusted R-squared (0.60 vs. 0.66).[72] Given our portfolios' large tracking errors and the distinct market environment our sample covers, we expect some differences between our results and those of

[70] Both define Size differently (total amount of debt outstanding and equity market capitalization), but Houweling and van Zundert (2017) show the performance of Size to be robust to these alternative definitions.
[71] Like ours, the sample period of Slimane et al. (2018) is more recent (2007 – 2017) compared to that of Houweling and van Zundert (2017) (1994 – 2015). With respect to a more detailed discussion concerning the performance of Low-risk over time see Section 6.3b.
[72] The authors' HY Low-risk portfolio produces a statistically significant alpha of ~2%, while that of ours is small and not statistically significant.

Houweling and van Zundert (2017). That they are nevertheless so similar, however, helps validate the quality of our data and lends credibility to our results beyond that to which our robustness tests entitle them. For example, limited data prevents us from examining alternative definitions of Value, Low-risk and Size. Additionally, while we construct credit rating and maturity-controlled factor portfolios (excluding Low-risk, which we intentionally construct on those two dimensions), we do not test whether sector-neutral factor portfolios would yield consistent results. However, since Houweling and van Zundert (2017) find their results to be robust to these tests, we propose it is likely that ours would be as well.

To conclude we examine the performances of long-only variants of our factor portfolios net of transaction costs (results of which are summarized in Table 4 below).[73] Given the well-known difficulties in shorting corporate bonds (Blitz et al., 2019), these portfolios more accurately reflect the achievable returns for the average investor (which we refer to as "investable" returns). In the case of IG Low-risk, IG Momentum, HY Value, HY Size and HY Credit, the alpha of our long-only portfolio is comparable (though noticeably lower) to its long-short counterpart. The remaining long-only factor portfolios, however, either significantly underperform their long-short versions, or fail to generate statistically significant alphas. The underperformance of long-only factors relative to long-short factors supports the criticism levied by Blitz et al. (2019) against researchers who rely on long-short factors without much regard for their practical implementation. The development of more realistic benchmarks is crucial, but given the nascent state of the literature concerning bond-specific factors, we follow the tradition of Fama and French (1993) and use long-short factor portfolios.[74] We will, however, report the "investable" returns of our top-performing portfolios at the end of our analyses (see Section 6.5).

6.2 Moving towards a multi-factor model

There are a number of complications that arise from our relatively crude construction of our factor-mimicking portfolios. Israel at al. (2018) explain that while long-short equity portfolios are largely beta-neutral, this is often not the case with long-short bond portfolios. For example, differences in average duration and credit ratings between a portfolio's long and short legs can result in significant embedded market exposure. To better isolate a particular factor, Israel et al. (2018) recommend a 5x5 bivariate sort, first on a beta proxy (in their case duration-times-spread) and subsequently on the characteristic measure of interest. This approach is similar to that of Fama and French (1993), who, to minimize the influence of Size on Value, sort stocks into three groups according to market capitalization (small, medium and large), and then into quintiles according to book-to-market ratios. As we explain in Section 5.2, we use bivariate sorts to construct Downside risk and Credit Quality. Our remaining factors, however, are simple univariate portfolios. To sort each of our factor portfolios according to every other factor, as a well as on a beta proxy, is not feasible, especially in our HY sample, which averages 250 bonds per month.[75] We therefore incorporate alternative procedures that help us

[73] We calculate transaction costs using estimates based on bond size, credit quality and maturity from Chen et al. (2007) and, where necessary, Harris (2015). See Section 5 for a more detailed explanation.

[74] The vast majority of researchers continue to rely on academic factors. In our case, doing so has the added benefit of helping us to eliminate embedded market exposure. If we had analyzed actual funds and not synthetic portfolios, it would be far more important that we develop an investable multifactor portfolio with which to assess fund performance.

[75] In regards to Low-risk, such a procedure would be counterproductive, as we intentionally construct the factor to capture high credit rating and short maturity effects.

to identify, and eliminate, unwanted market and factor exposure. These include the factor-spanning tests of Huberman and Kubel (1987), as well as the orthogonalization procedure of Elton et al. (1993).

As a preliminary analysis, we use simple bivariate sorts to create alternative factor portfolios (each of which incorporates one additional characteristic measure), and we examine the pairwise correlations between each of our factors (including our market portfolios and traditional bond and equity factors). With respect to the former, we find our results are mostly robust to controls on credit ratings, maturity and amount outstanding. Regarding the latter, we find that Size (IG and HY), Momentum (IG) and Credit Quality (IG and HY) exhibit either very low, or negative, correlations with the market portfolio and the other factors. Some of our more striking results are the very high, negative correlations between Low-risk and Downside risk in both IG (-0.95) and in HY (-0.83). This is not unexpected, however, as both factors are essentially beta proxies, and are highly correlated with their respective market portfolios. Somewhat more surprising are the relatively high correlations between Value and Low-risk, Downside risk and *DEF* (see Table 3). These correlations persist with Value portfolios that we have also sorted on maturity and credit ratings.[76] Our HY Momentum portfolio exhibits a high, negative correlation with the market portfolio (-0.75), but this is approximately in line with the strong negative correlation between the equity factors *UMD* and *RMRF* (-0.56). Lastly, consistent with Asness et al. (2015) and Israel et al. (2018), our factor-mimicking Value, Momentum and Size portfolios exhibit very low (and even negative) correlations with their equity market counterparts.

Many researchers, including Fama and French (2015, 2017), use factor-spanning tests to search for redundant factors. The test involves a regression of the returns of one factor against those of the remaining factors and a constant. A statistically insignificant intercept implies that the removal of the factor in question would not negatively impact the mean-variance-efficient tangency portfolio implied by the remaining factors. We conduct these tests on all of our factor-mimicking portfolios, our market portfolio and each of the conventional equity and bond market factors. Regressions on Market (IG, HY), Value (IG, HY), Momentum (HY), Size (HY) and Low-risk (IG) produce statistically significant intercepts at the 10% level. These tests give us some confidence that our factor-mimicking portfolios do not simply reflect market risk or exposure to other factors. Regarding traditional bond and equity market factors, we find that while *TERM* and *RMRF* generate statistically significant intercepts, *SMB*, *HML*, *UMD* and Option do not.[77] Although we remove the latter, we retain evidently redundant corporate bond factors.[78] Regarding Size (IG) and Low-risk (HY), we do this in order to align our IG and HY models. In the case of Credit Quality (IG, HY) and Downside risk (IG, HY), we find statistically significant intercepts in regressions that exclude Value and Low-risk. As we expect based on the correlation matrices above, Low-risk renders Downside risk redundant and vice-versa.[79] Additionally, Low-risk and Downside risk capture much of the variation in Credit Quality and Value respectively. Instead of discarding either Value/Low-risk or Downside risk/Credit Quality, we divide these factor pairs into two separate models.

---

[76] In fact, the correlations between the maturity-controlled Value portfolios and the aforementioned factors are even higher.

[77] *TERM* is somewhat surprising as our reliance on maturity-matched excess returns diminishes its importance (we find that it captures very little of the time-series variation in our portfolios' returns).

[78] The t-stats associated with these factors are between 1 and 1.6, which, at the very least, are markedly higher than those of SMB, HML, UMD and Option (all less than 0.5).

[79] Again, these results are intuitive given that Low-risk and Downside risk both represent beta proxies, and that we incorporate credit ratings in the construction of both Low-risk and Value.

Table 4: Performance of long-only factor portfolios

This table shows the performance statistics of Size, Value, 1-month Momentum, 6-month Momentum, Low-Risk, Downside, and Credit Rating factors for U.S. Investment Grade and U.S. High Yield corporate bonds over the period January 2015 to March 2021. See Section 4 for details regarding the construction of each factor portfolio. We calculate returns as the average of the portfolios constructed from month $t-1$ to month t. We measure corporate bond returns as excess returns versus maturity-matched Treasuries. Panel A reports statistics on the returns series for each factor. Panel B shows the CAPM-alpha and -beta with respect to the corporate bond market (DEF). Panel C shows alpha statistics with respect to a multi-factor model (RMRF, SMB, HML, MOM, TERM and DEF). We annualize all alphas and report the models' adjusted R^2. *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence levels, respectively, of two-sided tests on whether the alphas are different from 0 (t-tests with Newey-West standard errors).

	Investment Grade						
	Size	Value	Mom (1 mo)	Mom (6 mo)	Low Risk	Down-side Risk	Credit 2
Panel A: Returns Series Statistics							
mean	3.98%	2.97%	1.25%	1.79%	2.05%	4.07%	3.72%
volatility	4.76%	2.40%	6.40%	6.74%	2.00%	7.72%	4.95%
Sharpe ratio	0.84	1.24	0.20	0.27	1.02	0.53	0.75
skewness	-0.59	0.54	-0.70	-0.49	-0.13	-0.84	-1.47
kurtosis	1.51	1.89	2.96	2.69	2.67	2.38	7.08
Panel B: CAPM Results							
alpha	0.45%	1.54%*	-3.17%***	-2.86%**	0.56%**	-1.76%	-0.14%
t-value	0.779	1.730	-3.394	-2.544	2.082	-1.650	-0.350
beta	0.975	0.394	1.219	1.281	0.411	1.610	1.064
adj R^2	0.887	0.563	0.764	0.763	0.894	0.919	0.978
Panel C: Model 1 Results							
alpha	0.67%	1.23%*	-2.65%***	-2.35%**	0.49%*	-1.79%*	-0.04%
t-value	1.403	1.877	-2.769	-2.504	1.854	-1.733	-0.117
adj R^2	0.898	0.581	0.751	0.755	0.891	0.915	0.980

	High Yield						
	Size	Value	Mom (1 mo)	Mom (6 mo)	Low Risk	Down-side Risk	Credit 2
Panel A: Returns Series Statistics							
mean	5.43%	8.30%	8.63%	6.71%	2.30%	6.94%	3.63%
volatility	4.58%	3.66%	11.88%	8.54%	3.67%	14.61%	18.43%
Sharpe ratio	1.19	2.27	0.73	0.79	0.63	0.47	0.20
skewness	-2.03	0.18	-0.23	-1.08	-4.62	-1.55	-1.87
kurtosis	11.33	1.01	7.26	6.03	24.21	14.50	10.44
Panel B: CAPM Results							
alpha	2.68%***	6.41%***	1.81%	2.13%	-0.16%	-2.44%	-7.15%*
t-value	3.361	4.270	0.728	0.763	-0.320	-0.954	-1.715
beta	0.481	0.330	1.190	0.801	0.430	1.637	1.883
adj R^2	0.724	0.531	0.658	0.573	0.904	0.826	0.685
Panel C: Model 1 Results							
alpha	2.59%***	5.37%***	1.52%	2.61%	0.07%	-2.23%	-7.64%**
t-value	3.140	4.465	0.631	0.814	0.178	-0.881	-2.298
adj R^2	0.743	0.579	0.635	0.546	0.906	0.820	0.682

Model 2, which resembles the models featured in Houweling and van Zundert (2017) and Slimane et al. (2018), includes Value, Low-risk, Momentum and Size, while Model 3, which is analogous to the four-factor model of Bai et al. (2018), includes Downside risk, Credit quality, Size and Momentum.[80] Listed below are the three primary models we use to analyze the performance of ESG-sorted portfolios in the coming sections. We only use Model 4 in our final analysis in Section 6.5. Lastly, Table 5 summarizes the results of our factor-spanning tests on Model 2.

Model 1:

$$R_{p,t} - R_{f,t} = \alpha_i + \beta_{1,i} RMRF_t + \beta_{2,i} SMB_t + \beta_{3,i} HML_t + \beta_{4,i} TERM_t + \beta_{5,i} Option_t + \beta_{6,i} DEF_t + \varepsilon_{i,t}$$

Model 2:

$$R_{p,t} - R_{f,t} = \alpha_i + \beta_{1,i} DEF_t + \beta_{2,i} TERM_t + \beta_{3,i} Value_t + \beta_{4,i} Size_t + \beta_{5,i} MOM_t + \beta_{6,i} Low\ Risk_t + \varepsilon_{i,t}$$

Model 3:

$$R_{p,t} - R_{f,t} = \alpha_i + \beta_{1,i} DEF_t + \beta_{2,i} TERM_t + \beta_{3,i} Credit_t + \beta_{4,i} Size_t + \beta_{5,i} MOM_t + \beta_{6,i} Downside_t + \varepsilon_{i,t}$$

Model 4:

$$R_{p,t} - R_{f,t} = \alpha_i + \beta_{1,i} DEF_t + \beta_{2,i} TERM_t + \beta_{3,i} Value_t + \beta_{4,i} Size_t + \beta_{5,i} MOM_t + \beta_{6,i} Low\ Risk_t + \beta_i \sum Industry\ Dummies + \varepsilon_{i,t}$$

Factor-spanning tests help us to streamline our models, but embedded market and factor exposure still threaten to undermine the precision of our results. Of particular concern are the substantial relationships between the market portfolios and Value, Low-risk and Downside risk (and, to a lesser extent, between Low-risk and Value).[81] We implement the orthogonalization procedure of Elton et al. (1993), which consists of a regression of the returns of a given factor portfolio against those of the market portfolio (and/or of another relevant factor). We then combine the regression's intercept term and residual to create a new factor index, which exhibits a correlation with the market portfolio that is asymptotically zero. Remarkably, we find that the orthogonalization of Low-risk reduces its exposure to Value substantially, and the correlation between the two factors falls from -0.68 in IG and -0.61 in HY to -0.21 and -0.18 respectively. As a result, we do not feel it necessary to repeat the procedure to orthogonalize Value with respect to Low-risk.

[80] Bai et al. (2018) also include Momentum in their performance evaluation model. We substitute our Size portfolio for their Liquidity factor, which we are unable to reproduce due to data limitations. Importantly, our Momentum and Size portfolios appear to be insensitive to changes in Downside risk or Credit quality, and as such, we predict they will complement the other factors well.

[81] Which are evident in both the correlation matrices in Table 4 and the factor-spanning tests in Table 5.

Table 5: Factor spanning tests

This table displays the output from a set of factor spanning tests conducted on the factors in Model 2 for U.S. Investment Grade (IG) and U.S. High Yield (HY) corporate bonds over the period January 2015 to March 2021. See Section 5 for details regarding the construction of each factor portfolio. For each of the five factor portfolios, the table shows the coefficients of a regression of each factor against all others in Model 2. We annualize the model's Intercept term (alpha). *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence levels, respectively, of two-sided tests on whether the coefficients are different from 0 (*t*-tests with Newey-West standard errors).

	Investment Grade					*High Yield*				
	Market	Value	Size	Momentum (6mo)	Low-Risk	Market	Value	Size	Momentum (6mo)	Low-Risk
Intercept	1.89%***	2.13%	1.62%*	-5.03%**	1.02%*	3.83%**	4.24%*	1.67%*	8.85%*	1.46%
	(2.70)	(1.65)	(1.88)	(-2.22)	(1.73)	(2.28)	(1.89)	(1.87)	(1.83)	(1.08)
Market		-0.06	0.00	0.08	-0.71***		-0.13	0.09	-1.90***	-0.52***
		(-0.18)	(0.01)	(0.11)	(-18.03)		(-0.53)	(0.37)	(-4.15)	(-9.11)
Value	-0.01		0.10	0.13	-0.11***	-0.06		0.03	-0.24	-0.24***
	(-0.18)		(1.45)	(0.37)	(-2.67)	(-0.52)		(0.26)	(-0.65)	(-2.76)
Size	0.00	0.40**		1.11***	0.01	0.04	0.03		0.38	0.05
	(0.01)	(2.02)		(2.91)	(0.09)	(0.36)	(0.26)		(1.61)	(0.68)
Momentum (6mo)	0.01	0.06	0.13***		0.00	-0.20**	-0.06	0.09		-0.04
	(0.11)	(0.35)	(4.32)		(0.12)	(-2.60)	(-0.67)	(1.51)		(-0.79)
Low-Risk	-1.09***	-1.05**	0.02	0.08		-0.77***	-0.86***	0.17	-0.54	
	(-13.48)	(-2.57)	(0.09)	(0.12)		(-5.42)	(-3.22)	(0.63)	(-0.83)	
Equity Market	-0.02	-0.07**	-0.03	-0.01	-0.02	-0.05	0.01	-0.05	-0.07	-0.01
	(-1.20)	(-2.14)	(-1.63)	(-0.11)	(-1.36)	(-1.19)	(0.26)	(-0.95)	(-0.62)	(-0.45)
Equity MOM	-0.05**	0.07*	-0.06***	-0.02	-0.03	-0.08**	-0.19***	-0.07	-0.15	-0.12***
	(-2.58)	(1.85)	(-3.03)	(-0.42)	(-1.38)	(-2.04)	(-2.87)	(-1.28)	(-1.14)	(-3.35)
HML	-0.05*	-0.01	-0.02	0.08	-0.03	-0.08	-0.12	0.01	-0.01	-0.09*
	(-1.88)	(-0.17)	(-1.09)	(1.00)	(-1.17)	(-1.30)	(-1.22)	(0.17)	(-0.04)	(-1.69)
SMB	0.00	0.12***	-0.01	-0.19**	0.02	-0.08	-0.05	-0.03	-0.20	-0.09**
	(0.20)	(3.04)	(-0.27)	(-2.08)	(0.94)	(-1.23)	(-0.71)	(-0.25)	(-0.87)	(-2.08)
TERM	-0.00	-0.34*	0.02	0.21	-0.02	-0.16	0.09	-0.00	-0.19	0.02
	(-0.02)	(-1.77)	(0.25)	(0.75)	(-0.29)	(-1.27)	(0.27)	(-0.00)	(-0.46)	(0.16)
adjusted R^2	0.857	0.486	0.240	0.179	0.868	0.738	0.323	0.003	0.552	0.683

6.3 Examination of ESG-sorted portfolios

6.3.1 Quintile Portfolio Analysis

We use quintile analysis to begin our examination of the relationship between ESG variables and excess returns. This allows us to simultaneously survey differences in the financial returns of various ESG-sorted portfolios together with the firm- and bond-level characteristics that may help to explain them (Bennani et al. (2019), Dorfleitner et al. (2020)). In conjunction with evidence from the literature that IG and HY bonds belong to distinct markets (Chen et al., 2014, Houweling and van Zundert, 2017), the differences we observe in the returns to IG and HY factor portfolios lead us to separate the two throughout our analyses. We also follow Brammer and Millington (2012) in creating separate quintile portfolios for each ESG pillar score (which include Environmental, Social, Governance and Combined scores). The authors argue that researchers who fail to do so obscure trends specific to individual pillars.[82] Table 6 displays performance statistics and average firm and bond characteristics for each quintile portfolio. Though not summarized in the table, we also analyze ESG portfolios that we construct with alternative cut-off rates of 10% and 50%. We reference the results of these analyses only when they contrast with those of the quintile portfolios. Like Pereira (2019), we find that our results are mostly robust to changes in cut-off rates.

The most pronounced relationships across IG quintiles appear to be those between average ESG scores and both bond size and credit quality. Average bond size (measured as amount notional outstanding) decreases almost monotonically from top quintile portfolios to bottom portfolios, such that, among Combined, Environmental and Social portfolios, the average bond size of the top quintile is 93%, 75% and 115% larger than that of the respective bottom quintile. There is little variation in bond size between Governance portfolios, but bonds in the top quintile are still on average 18% larger than those in the bottom. We find a similar trend with respect to average credit quality, which deteriorates from quintile one to quintile five in Combined, Environmental and Social portfolios.[83] Credit quality varies only slightly across Governance portfolios, which is curiously consistent with Polbennikov et al. (2015), who even detect a slight negative correlation between governance and credit ratings.[84] Lastly, while there is no obvious relationship between average maturity and ESG ratings in Combined and Environmental quintiles, we see strong, opposing relationships in Governance and Social portfolios. The bonds that comprise the top-rated Governance portfolio have, on average, the longest maturities (11.64 years), which are nearly two standard deviations greater than those of bonds in the bottom quintile portfolio (7.96 years).[85] Among Social quintiles, however, average maturity steadily increases from the top portfolio (8.37 years) to the bottom (10.05).

With the exception of those concerning our Governance portfolios, our results in IG find support in the existing literature. To the extent that bond size correlates with firm size, the relationship we observe between average ESG scores and amount outstanding aligns with the recent findings of Drempetic et al. (2019). The authors show that, after controlling for a host of firm-specific factors,

[82] In fairness, some of the studies to which the authors refer simply use ESG ratings to proxy for intangibles like social capital, reputation and trust (more recent examples include Lins et al. (2017) and Amiraslani et al. (2019)).
[83] That top quintile portfolios display higher credit ratings translates to lower numerical scores.
[84] Like the authors, we have no economic rationale for the lack of a relationship between governance scores and credit ratings.
[85] Changing the cut-off rate to 10% does not materially increase or decrease the average maturity of the top and bottom Governance portfolios (which become 11.92 and 8.01 years respectively).

large companies receive higher ESG ratings on average.[86] They attribute this to both slack resource theory, which suggests that larger and more profitable companies have more resources to invest in ESG activities, and to legitimacy theory, which assumes that such companies face greater social pressure to disclose sustainability information (Gavana et al., 2017 and Drempetic et al. 2019).[87] Concerning credit ratings, our results fit within a long stream of research that demonstrate an association between strong ESG performance and both lower credit risks and higher credit ratings (see literature reviews in Oikonomou et al. (2014), Salvi et al. (2019)). Lastly, the relationship we see between Social scores and maturity is consistent with Benlemlih (2014), who shows that a firm's CSR performance (especially on Social dimensions) relates negatively to the average maturity of its bonds.[88] While no study of which we are aware discovers a comparable pattern in regards to corporate governance and maturity, we suggest an explanation along the lines of Shi and Sun (2015), who argue that shareholders and creditors of firms with strong governance mechanisms demand fewer channels through which to discipline management.[89]

We do not detect many of the same trends in the cross-sectional characteristics of HY quintile portfolios. The most notable similarity concerns the negative relationship between a portfolio's average maturity and its Governance score. However, while maturity steadily decreases from the top portfolio to the bottom portfolio, the gap between the two is much narrower in HY (1.25 years compared to 3.68 years in IG). Regarding credit quality, top quintiles in HY exhibit better (albeit slightly) average credit ratings compared to bottom quintiles, but there is no consistent trend across the intermediate portfolios. The most distinct relationship in IG, that between a portfolio's average bond size and its ESG score, is entirely absent in HY. That this relationship breaks down, we propose, is consistent with trade-off theory (Aupperle, 1985). Although a large amount of debt outstanding might indicate a large company for which spending on ESG is comparably cheap (Dorfleitner et al., 2020), it may also indicate greater leverage and higher total debt costs, which, we suggest, is more likely to be the case in HY.[90] A HY borrower is likely to have fewer slack resources, as well as a higher required rate of return, both of which make investments in ESG less attractive.

Next we examine the average excess returns and risk-adjusted performance of ESG-sorted portfolios. Across IG quintile portfolios there appear to be no clear relationships between average ESG ratings and annual excess returns. The top and bottom quintiles with respect to Combined and Environmental scores generate approximately the same average returns (3,72% and 3,6%), while the top Social portfolio marginally underperforms the bottom Social portfolio. The top two Governance quintiles meaningfully outperform the bottom two (average excess returns of 4.10% vs. 3.36%), but the return differential is less than one standard deviation. A pattern emerges, however, when we look at a portfolio's risk-adjusted performance measured by the Sharpe ratio. Risk-adjusted returns mostly increase from the top IG quintile portfolios to the bottom portfolios. While certain quintiles disrupt

[86] The relationship between bond size and firm size diminishes in a sample that includes HY firms, as the total amount of bonds outstanding also relates to credit quality (since total debt outstanding is an important determinant of a firm's credit rating).
[87] Stanwick and Stanwick (1998) also show company size to be an important determinant in sustainability performance. Drempetic et al. (2019) emphasize the relationship between firm size and ESG ratings, which they note does not necessarily represent ESG performance.
[88] Which they explain as a display of trust as it exposes a firm to a greater degree of debt discipline.
[89] Short-term debt disciplines borrowers by decreasing incentives for suboptimal risk taking that can arise from conflicts of interest between lenders and borrowers (Myers, 1977, Chen et al., 2018).
[90] To make more qualified comments would require us to conduct a more focused analysis of the relationship between bond size, company size, leverage, and ESG scores in both IG and HY. Since we do not have data on company size (or at least insufficient coverage), we save such analyses for future research.

this trend, notably Environmental portfolio 2 and Governance portfolio 4, in all cases, bottom portfolios boast better risk-adjusted returns than top portfolios.[91]

Top quintiles in HY also appear to underperform portfolios with lower average ESG ratings. In most cases, this disparity in financial performance is more substantial in HY. For example, the difference in the Sharpe ratios of the top and bottom Combined quintiles is a full standard deviation larger in HY (-0.874 vs. a difference of -0.142 in IG). Unlike in IG, however, it appears that middle HY portfolios tend to outperform their respective bottom portfolios. In regards to Combined, Environmental and Governance quintiles, portfolios comprised of bonds with the top 20% - 60% ESG scores exhibit distinctly higher average excess returns and Sharpe ratios compared to those containing the remaining bonds. The top Environmental quintile produces a measly Sharpe ratio of 0.47, which is nearly one standard deviation below the mean, and almost two standard deviations below the Sharpe ratio of the second portfolio (0.88). The average Sharpe ratio of the middle three Governance portfolios is also considerably higher than that of the top and bottom portfolios (0.83 vs. 0.62). Due to the strong performances of the middle quintiles, a portfolio of the 50% highest-rated bonds slightly outperforms that of the 50% worst-rated bonds.[92] Social portfolios seem to behave differently than the others. While the top two Social quintiles generate higher average excess returns, Social quintiles 3 and 4 exhibit higher Sharpe ratios. The magnitude of these differences, however, is smaller among Social portfolios than it is among Combined, Environmental and Governance portfolios. Interestingly, the bottom Social quintile is the worst performer on both an absolute and a risk-adjusted basis, making it the only bottom quintile portfolio across HY and IG to underperform all of the quintiles above it.

That top quintiles underperform lower quintiles in both IG and HY supports out prediction that investments into bonds from issuers with high ESG ratings do not generate above market returns. These results are mostly consistent with those of Dorfleitner et al. (2020), who find that, across equally-weighted quintile portfolios of equities, those with the lowest average ESG ratings perform best. The authors ascribe this phenomenon to the market's asymmetric treatment of the ESG practices of large and small companies, an explanation that, at least among IG issuers, may also apply in our case with respect to bond size. Our results could also fit with a risk-based explanation like that of Pastor and Stambaugh (2020), such that low-rated bonds earn a premium for exposure to a systematic ESG risk factor.[93] In any case, average excess returns and Sharpe ratios are insufficient to assess a portfolio's financial performance. Additionally, as our examination of ESG quintiles lays bare, we need to control for factors that may influence both ESG ratings and returns. In the following sections, we use our multifactor models, as well as alternative methods of portfolio formation, to control for the potentially confounding effects of prominent bond characteristics and risk factors.

[91] One minor exception concerns the top and bottom Governance portfolios formed at a 50% cut-off, in which the top portfolio exhibits both higher excess returns and a higher Sharpe ratio (though the differences of 0.42% and 0.085 respectively are not significant). However, most of these differences are rather slight, and only that between top and bottom Environmental portfolios exceeds one standard deviation.

[92] At a 50% cut-off, the top portfolio generates average excess returns of 5.85% and a Sharpe ratio of 0.735 compared to 5.37% and 0.692 of the bottom portfolio respectively.

[93] Though again this appears to apply more so to IG bonds, as it is only in IG quintiles that performance seems to increase incrementally from the top portfolio to the bottom portfolio.

6.3.2 Multivariate regression analysis

In the following section we examine the extent to which our three multifactor models explain the returns to our ESG-sorted quintile portfolios. The first of these models includes only traditional bond and equity market factors, and closely resembles the pricing models of Houweling and van Zundert (2017), Israel et al. (2018), Bai et al. (2018) and Bali et al. (2019) (see Model 1, 6.2a). Model 2 includes our Size, Momentum, Value and orthogonalized Low-Risk factors, as well as *DEF*, TERM and EQUITY, while Model 3 replaces Value and Low-Risk with our orthogonalized Downside-Risk and Credit Quality factors. Since alphas are notoriously model-specific, we test each of our portfolios against all three models. This includes our quintile portfolios as well as our long-short portfolios, which go long the top quintile and short the bottom quintile with respect to each ESG rating category. Table 7 displays the annualized alphas, t-statistics, and adjusted R-squared values from these regressions. As in Section 6.1, we use the approach of Newey and West (1987) to calculate heteroskedastic and autocorrelation robust standard errors. We also check for multicollinearity, but following the orthogonalization procedure in Section 6.2a, we never detect variance inflation factors (VIFs) above 2.5. In Table 8, we report the regression coefficients from Model 2.

Lewellen et al. (2010) propose that researchers should evaluate a new multifactor model in terms of its ability to explain the time-series variation in the returns of a set of test portfolios. To create these portfolios, the authors suggest that researchers group securities according to characteristics that are unrelated to the factor-mimicking portfolios which comprise the model.[94] For example, Bai et al. (2018) form twenty-five portfolios based on bond size and maturity with which to assess the effectiveness of their four-factor model (containing Downside-risk, Credit Quality, Momentum and Liquidity factors) relative to that of a traditional model (which is comparable to Model 1). We argue that the nascent state of the literature concerning bond-specific factors makes this analysis especially important. We also suggest that our ESG-sorted portfolios are rather ideally suited to act as test portfolios, and thus before we delve into our main analysis, we briefly comment on the apparent efficacy of our models.

Models 2 and 3 significantly outperform Model 1 in terms of explanatory power measured by the adjusted R-squared. This is particularly the case with some of the long-short portfolios that are mostly market-neutral, such as the long-short Social portfolio in IG (adjusted R-squared increases from 0 to roughly 0.3) and the long-short Governance portfolio in HY (adjusted R-squared increases from 0.02 to approximately 0.5). The average adjusted R-squared across all long-only portfolios is 0.892 for Model 1, and 0.932 and 0.924 for Model 2 and Model 3 respectively. In general, the R-squared values of our multifactor models are markedly higher than those of Bai et al. (2018), which we attribute to our replacement of the authors' *DEF* factor with an equally-weighted portfolio of all bonds in our sample, as well as our construction of separate factor-mimicking portfolios in IG and HY.[95] Perhaps most importantly, and as we discuss in more detail below, there are instances where large and statistically significant alphas based on Model 1 (such as that of Combined portfolio 3 in IG), disappear entirely in Models 2 and 3. While we hope our use of these models helps us to minimize misspecification biases, we are wary of potential problems that accompany relatively untested multifactor models. For

[94] As intuitive as this would appear to be, many researchers create test portfolios based on characteristics that are represented in their factor model (i.e. they test a model containing a Size factor using size-sorted quintile portfolios) (Fama and French, 2017).

[95] The authors use the traditional conceptualization of *DEF* as the returns to an aggregate bond index. Our use of a sample-specific market portfolio is consistent with the most recent literature (Houweling and van Zundert, 2017, Bektic et al., 2018, Hoepner and Nilsson, 2020).

example, issues related to accidental data mining or overfitting could inflate our models' R-squared values artificially.[96] We note, however, that our results fit within those from studies that employ comparable models (Bali et al., 2019, Hoepner and Nilsson, 2015, 2020).[97] Additionally, we always display the alphas from Model 1 (which includes only conventional factors) along with those of our new models. We propose that our results indicate a rather low likelihood that our new models suffer from some unidentified problem that does not also afflict the traditional model.

The results of our regressions on IG portfolios seem to confirm many of the trends we observe in the cross-section. Most notably, we find that lower quintile portfolios outperform top portfolios across each ESG rating category. In accounting for prominent factor exposures, however, these regressions also reveal some distinctions that we could not detect based on bond characteristics and Sharpe ratios alone. Regarding Combined quintile portfolios, for example, Table 6 suggests that quintiles 3 and 4 perform best. According to our multivariate regressions, these portfolios produce small and statistically insignificant intercepts, while portfolio 5 generates a relatively robust alpha of 0.7%. Similarly, among portfolios formed on individual pillar scores, we no longer see a clear, negative relationship between average ESG ratings and financial performance. Top portfolios produce weaker returns and Sharpe ratios compared to lower-rated portfolios, but, according to Model 2, exhibit rather neutral alphas on par with those of middle quintiles.[98] The only portfolios associated with consistently positive and statistically significant alphas are the bottom quintile portfolios, which, in all cases, are the best performing quintile portfolios (alphas range from 0.5% to 1%). These results appear to be the most consistent with a demand-based effect, like the shunned-stock effect, where the negative influence of ESG on returns is confined to the lowest rated issuers (Derwall et al., 2011, Pereira, 2019). Derwall et al. (2011) outline how researchers can search for the shunned-stock effect in the time series of average returns, which we analyze in section 6.4.[99]

The results in Table 7 also recontextualize our previous observations with respect to ESG ratings and HY bond returns. For example, HY Governance portfolio 3, which along with Governance portfolio 2 displays the strongest average excess returns and Sharpe ratios, produces a small and statistically insignificant alpha (even one that is slightly negative based on Model 2). On the other hand, Governance portfolio 2 produces an economically and statistically robust alpha of ~1.6%. We see a similar phenomenon regarding Environmental portfolios 2 and 3, which also show comparable average excess returns and Sharpe ratios. Environmental portfolio 2 generates a statistically significant intercept of 1.3%, while that of portfolio 3 is small and statistically insignificant.

---

[96] It is also possible that our factors correlate with some unobservable latent risk factor that could serve to artificially inflate intercept values (especially since the determinants of corporate bond factors are still not very well understood).

[97] In an attempt to create the most effective multifactor model to explain corporate bond returns, Hoepner and Nilsson (2018) arrive at R-squared values in excess of 0.98. Given the trial-and-error nature of their development process, we are concerned that they fit their models a little too closely to the data. [[[I can see that this might be a little too presumptuous of me, I recognize my limitations and that I am certainly not in a position to criticize another paper in this way]]].

[98] The top Combined portfolio generates a small and (weakly) statistically significant alpha of -0.3% based on Model 1, but this negative alpha disappears entirely in Models 2 and 3.

[99] Given the time-series trends we observe in Section 6.4, as well as the fact that abnormal returns are largely confined to bonds with the worst ratings, we suggest our results in IG fit more neatly with the shunned-stock effect than with the presence of an ESG risk premium (Pastor et al., 2020), or, as Hong and Kacperczyk (2009) so eloquently put it, a payment for being "bad".

Table 6: Characteristics of ESG-sorted long-only portfolios

This table shows the mean value of various bond characteristics in our ESG-sorted portfolios for U.S. Investment Grade (IG) and U.S. High Yield (HY) corporate bonds over the period January 2015 to March 2021. We construct the portfolios by taking long positions in bonds per quintile of their issuers' Thomson Reuters (TR) ESG Combined score, Environment score, Social score, and Government score. We measure Credit Rating as the average across S&P, Moody's and Fitch ratings, Amt Out as the amount of debt outstanding, and corporate bond returns as annualized excess returns versus maturity-matched treasury bonds.

	Investment Grade								*High Yield*							
	ESG Score	Credit Rating	Amt Out	Yrs to Maturity	Duration	Excess Returns	Volatility	Sharpe ratio	ESG Score	Credit Rating	Amt Out	Yrs to Maturity	Duration	Excess Returns	Volatility	Sharpe ratio
ESG																
top 20%	81.21	7.62	717.4	9.49	6.58	3.78%	5.15%	0.734	77.72	10.99	531.1	8.45	6.09	4.91%	8.57%	0.573
20-40%	71.86	7.82	644.1	8.77	6.21	3.61%	4.61%	0.783	62.88	11.40	512.7	9.16	6.42	6.87%	8.42%	0.816
40-60%	64.67	7.95	517.1	8.41	6.23	3.60%	4.27%	0.843	52.66	12.17	514.8	8.63	6.07	6.41%	8.37%	0.766
60-80%	54.86	7.94	496.5	9.03	6.15	3.91%	4.62%	0.847	42.73	11.54	447.0	8.01	5.70	4.96%	7.48%	0.663
bot 80%	35.92	8.20	372.6	8.84	6.13	3.66%	4.46%	0.822	25.99	11.77	533.1	7.18	5.35	5.02%	7.01%	0.716
Environment																
top 20%	76.91	7.57	703.5	8.97	6.42	3.59%	5.10%	0.703	74.97	11.17	575.7	8.68	6.20	4.32%	9.26%	0.466
20-40%	70.67	7.80	567.8	8.77	6.27	3.87%	4.38%	0.882	61.18	11.18	508.6	9.28	6.43	6.74%	7.67%	0.878
40-60%	65.81	7.99	548.0	9.37	6.45	3.77%	4.92%	0.765	52.61	12.03	495.6	8.30	5.86	6.23%	7.66%	0.813
60-80%	55.48	7.99	531.3	9.57	6.38	3.70%	4.62%	0.801	42.63	11.88	378.9	7.28	5.34	5.81%	7.88%	0.737
bot 80%	39.29	8.17	391.4	7.86	5.78	3.66%	4.08%	0.898	31.46	11.60	573.0	8.08	5.90	5.18%	7.21%	0.718
Social																
top 20%	78.89	7.47	768.8	8.37	6.29	3.64%	4.63%	0.786	75.74	11.12	545.4	8.42	6.11	6.28%	9.16%	0.685
20-40%	70.20	7.86	571.2	8.63	6.11	3.45%	4.36%	0.792	60.84	11.45	473.3	8.86	6.28	6.27%	9.13%	0.688
40-60%	63.81	7.86	544.6	8.98	6.29	3.57%	4.69%	0.761	51.21	11.79	480.8	8.46	5.96	5.33%	7.55%	0.706
60-80%	56.30	8.06	477.0	8.49	6.16	4.08%	4.85%	0.841	42.71	11.51	507.4	7.82	5.63	5.65%	6.72%	0.840
bot 80%	39.11	8.28	381.2	10.05	6.44	3.84%	4.60%	0.833	31.81	11.99	523.6	8.04	5.74	4.59%	7.48%	0.614
Government																
top 20%	72.99	7.92	617.9	11.64	6.88	4.06%	5.21%	0.778	65.84	11.24	420.5	8.94	6.21	5.39%	8.31%	0.649
20-40%	66.70	8.10	556.8	8.91	6.49	4.16%	4.95%	0.841	59.56	11.64	537.1	8.58	6.14	7.47%	8.73%	0.855
40-60%	63.28	7.85	561.5	7.88	5.96	3.61%	4.25%	0.851	53.24	11.95	528.0	8.29	5.87	5.90%	7.27%	0.811
60-80%	55.23	7.84	518.0	8.16	6.02	3.20%	4.63%	0.692	50.24	11.51	492.9	8.09	5.87	4.50%	8.47%	0.531
bot 80%	49.83	7.83	482.3	7.96	5.96	3.57%	4.14%	0.862	33.80	11.56	551.7	7.69	5.63	4.95%	7.22%	0.686

The alpha of Environmental portfolio 2 stands in even starker contrast to that of portfolio 1 (-2.5%). In fact, that second portfolios outperform top portfolios is the most notable similarity between HY and IG.[100] This trend, however, is much more pronounced in HY, and, among portfolios formed on Combined, Environmental and Governance scores, the second quintiles are the best performers while the top quintiles are the worst.[101] Lastly, though all long-short portfolios in IG and HY generate negative alphas, they appear to do so for different reasons. In IG, these alphas reflect the strong returns of bottom quintile portfolios, whereas in HY they are attributable to the poor performances of the top portfolios.[102]

The overperformance of middle quintile portfolios relative to top and bottom portfolios in HY is difficult to reconcile with the predictions of either the shunned-stock or errors-in-expectations hypotheses, or with an explanation regarding an ESG risk premium (Pastor et al., 2020). However, although our findings contrast with many traditional models of the CSP-CFP relationship (Kempf and Ostoff, 2007, Derwall and Koedjik, 2009), they align well with the very recent theoretical frameworks of Barth et al. (2019) and Hoepner and Nilsson (2020). Barth et al. (2019) explain that, while poor ESG practices can increase credit risks, the marginal returns to incremental spending on ESG are sharply decreasing. Excessive spending on ESG, reflected in the very highest ESG scores, destroys value and, via Merton (1974), increases credit spreads. Hoepner and Nilsson (2020) build on Barney's (2019) stakeholder-based resource theory (SBRT) and suggest that, in many cases, the variable payoffs to investments in ESG inherently benefit residual claimants (shareholders) more so than they do fixed claimants (bondholders). The authors propose that the risks of asset substitution rise alongside ESG spending, causing bondholders to penalize issuers with high ESG scores. Bonds from low-rated firms, however, underperform as well, as poor ESG practices signal heightened risks of both credit events and of agency frictions between fixed and residual claimants (Bauer and Hann, 2010).

The existing literature also provides potential reasons as to why our results in HY appear to fit more neatly with the explanations above than do those in IG. Dorfleitner et al. (2020) reference slack resource theory and trade-off theory to explain why the market apparently punishes high ESG scores in small companies but not in large companies. As it is for the former, spending on ESG is more costly for HY issuers, which face a higher required rate of return and which are less likely to have spare capital to allocate towards investment in ESG. Such firms are not only more exposed to risks associated with greater ESG spending, but also with those that arise from poor sustainability practices. According to Amiraslani et al. (2019), a firm's mismanagement of CSR reduces the level of trust it has with its stakeholders, which, especially during periods when overall trust is low (i.e. financial crises), manifests in a higher likelihood of agency conflicts and even bankruptcy. Since companies closer to default are more prone to these risks, creditors of HY firms are more sensitive to a firm's ESG record. Taken together, these two studies suggest that bondholders of HY firms are more likely than those of IG firms to penalize both poor ESG practices and excessive ESG spending. This could explain why top and bottom quintiles in HY (but not in IG) underperform middle quintile portfolios.

Differences between our results from Models 1, 2 and 3 may provide additional insight into the trends we have discussed so far. The positive alphas associated with most of our bottom quintile

[100] In IG this is particularly the case with Environmental portfolios (alphas of 0 and 0.8% in top and second quintile portfolios respectively).
[101] Although the top Social portfolio also underperforms its peers (and is the only Social quintile to produce a statistically significant intercept (~-1%)), the second portfolio generates an alpha that is effectively 0.
[102] Just as the bottom IG Governance portfolio generates the smallest overperformance, so too does the top HY Governance portfolio exhibit the smallest underperformance (alpha is not statistically significant).

portfolios in IG decrease in both size and statistical significance as we move from Model 1 to Models 2 and 3. We similarly see some moderation in the negative alphas of our top portfolios, but these are small and statistically insignificant to begin with. The regression coefficients summarized in Table 8 help us to understand these results. In line with our observation of an association between bond size and ESG ratings in Table 6, we find that, across each top and bottom quintile portfolio in IG, factor loadings on Size are among the most consistently robust and statistically significant (only the coefficient on Market is stronger). All of the top portfolios produce large, negative coefficients on Size, while the bottom portfolios exhibit positive loadings of a similar magnitude.[103] Given that we detect a Size premium in IG (albeit a weak one), it makes sense that alphas decrease after we account for a portfolio's exposure to the Size factor. It is also noteworthy that top and bottom portfolios display statistically significant loadings on Momentum, which aligns with the results of Alessandrini (2019) regarding ESG-sorted equity portfolios. The author ascribes his result to the relatively larger exposure of high-rated equities to demand-based pressures, as well as to the overall increase in the volume of funds dedicated to SRI investing. We suggest that this explanation applies not just to securities with high ESG ratings, but to those with extreme values in either direction.

In contrast to our results in IG, the negative alphas associated with top HY portfolios increase as we move from Model 1 to Model 2. One possible reason for this, we suggest, concerns the degree to which Value accounts for the returns of these portfolios. Table 7 reveals that coefficients on the Value factor are large and positive among top quintile portfolios and negative among bottom portfolios (Social portfolios prove once again the exception). Crudely translated, our Value factor measures bond "cheapness" (Correira et al., 2012), at least relative to a bond's maturity and credit rating. In other words, in HY we find that highly rated portfolios are more exposed to the returns of cheap bonds than are portfolios containing bonds from issuers with lower ratings. Exposure to the Value factor, however, does not decrease monotonically from the top quintile to the bottom. For example, second quintile portfolios based on Combined and Environmental scores exhibit negative coefficients that are both larger, and more statistically significant, than those of their respective bottom portfolios. In general, it appears that middle portfolios tend to exhibit the highest exposure to expensive bonds. While consistent with the cross-sectional patterns we see in risk-adjusted performance, these results seem to oppose those in the existing literature, which presents evidence of a premium on high ESG bonds (Oikonomou 2014, Zerbib, 2018 Halling 2020).[104] We do, however, see some indication of this in IG, where all top (bottom) portfolios display negative (positive) coefficients on Value (though only those from Governance quintiles are statistically significant). These peculiar results further highlight a divergence in bondholders' perceptions of ESG in HY and IG, and support (however loosely) our earlier interpretation that bondholders are more likely penalize HY firms for high ESG scores (or, more accurately, the high levels of ESG investment that these scores represent).

[103] Coefficients related to the Governance portfolios are much smaller. This is consistent with the limited variability we see in average bond size across these portfolios (see Table 5 and our discussion in Section 6.1).

[104] It is not, however, an apples-to-apples comparison, as the literature defines the ESG premium in terms of credit spreads, whilst we discuss coefficients on our Value factor.

Table 7a: ESG multi-factor performance statistics (Combined and Environmental)

This table shows the performance statistics of quintile portfolios sorted on Thomson Reuters ESG and Environment scores for U.S. Investment Grade (IG) and U.S. High Yield (HY) corporate bonds over the period January 2015 to March 2021. In two sets of three columns, we show the alphas (with adjusted t-statistics) and adjusted R-squared values from regressions using Model 1, Model 2 and Model 3, which are described in Section 6.2a. We annualize the model's Intercept term (alpha). *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence levels, respectively, of two-sided tests on whether the coefficients are different from 0.

	Investment Grade						*High Yield*					
	Model 1		Model 2		Model 3		Model 1		Model 2		Model 3	
	Alpha (t-val)	adj R^2	Alpha (t-val)	adj R^2	Alpha (t-val)	adj R^2	Alpha (t-val)	adj R^2	Alpha (t-val)	adj R^2	Alpha (t-val)	adj R^2
ESG												
top 20%	-0.32%	0,966	-0.08%	0,984	-0.13%	0,985	-1.53%	0,929	-2.31%**	0,948	-2.40%**	0,942
	(-1.16)		(-0.38)		(-0.58)		(-1.47)		(-2.48)		(-2.33)	
20-40%	-0.21%	0,963	-0.42%	0,980	-0.45%	0,976	1.74%**	0,917	1.10%	0,953	1.30%	0,956
	(-0.92)		(-1.52)		(-1.39)		(2.21)		(1.10)		(1.44)	
40-60%	0.51%	0,937	0.31%	0,956	0.31%**	0,967	0.57%	0,909	0.54%	0,924	0.93%	0,920
	(1.48)		(0.87)		(2.45)		(0.78)		(0.57)		(0.86)	
60-80%	0.30%	0,960	0.02%	0,980	0.02%	0,978	-0.58%	0,934	-0.67%	0,943	-0.74%	0,942
	(1.33)		(0.09)		(0.08)		(-0.68)		(-0.88)		(-0.93)	
bot 80%	0.86%***	0,949	0.76%*	0,979	0.66%**	0,980	0.65%	0,952	0.91%	0,961	0.15%	0,967
	(2.82)		(1.92)		(2.49)		(0.89)		(1.27)		(0.26)	
Long/Short	-1.18%**	0,222	-0.84%**	0,583	-0.79%**	0,589	-2.18%	0,188	-3.22%**	0,351	-2.54%*	0,312
	(-2.60)		(-2.52)		(-2.23)		(-1.42)		(-2.24)		(-1.76)	
Environment												
top 20%	-0.40%	0,972	-0.26%	0,982	-0.39%	0,985	-2.43%**	0,933	-2.54%***	0,946	-3.17%***	0,938
	(-1.37)		(-0.95)		(-1.20)		(-2.35)		(-3.32)		(-3.17)	
20-40%	0.93%*	0,949	0.79%**	0,963	0.89***	0,971	1.68%**	0,936	1.55%*	0,949	1.74%	0,965
	(1.91)		(2.13)		(3.41)		(2.31)		(1.80)		(1.20)	
40-60%	-0.12%	0,970	-0.57%*	0,976	-0.54%	0,973	0.69%	0,926	0.60%	0,938	0.96%	0,939
	(-0.40)		(-1.70)		(-1.36)		(0.83)		(0.76)		(1.16)	
60-80%	-0.04%	0,972	0.03%	0,980	0.09%	0,980	0.29%	0,902	0.08%	0,922	0.57%	0,916
	(-0.14)		(0.12)		(0.31)		(0.41)		(0.10)		(0.77)	
bot 80%	0.94%***	0,960	0.67%***	0,978	0.71%***	0,977	0.75%	0,937	0.32%	0,948	0.01%	0,949
	(2.80)		(3.19)		(3.33)		(0.76)		(0.87)		(0.01)	
Long/Short	-1.34%***	0,314	-0.93%**	0,624	-1.16%***	0,641	-3.18%*	0,211	-2.86%***	0,403	-3.18%**	0,319
	(-2.67)		(-2.34)		(-2.79)		(-1.96)		(-2.70)		(-2.17)	

Table 7b: ESG multi-factor performance statistics (Social and Governance)

This table shows the performance statistics of quintile portfolios sorted on Thomson Reuters ESG and Environment scores for U.S. Investment Grade (IG) and U.S. High Yield (HY) corporate bonds over the period January 2015 to March 2021. In two sets of three columns, we show the alphas (with adjusted t-statistics) and adjusted R-squared values from regressions using Model 1, Model 2 and Model 3, which are described in Section 6.2a. We annualize the model's Intercept term (alpha). *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence levels, respectively, of two-sided tests on whether the coefficients are different from 0.

	Investment Grade						*High Yield*					
	Model 1		Model 2		Model 3		Model 1		Model 2		Model 3	
	Alpha (t-val)	adj R^2	Alpha (t-val)	adj R^2	Alpha (t-val)	adj R^2	Alpha (t-val)	adj R^2	Alpha (t-val)	adj R^2	Alpha (t-val)	adj R^2
Social												
top 20%	-0.00%	0,972	-0.15%	0,980	-0.05%	0,981	-1.24%*	0,947	-0.93%*	0,962	-1.37%**	0,962
	(-0.01)		(-0.61)		(-0.24)		(-1.81)		(-1.79)		(-2.03)	
20-40%	0.07%	0,939	0.09%	0,950	0.41%	0,961	0.29%	0,909	0.26%	0,920	-0.20%	0,937
	(0.18)		(0.25)		(1.22)		(0.26)		(0.21)		(-0.19)	
40-60%	-0.18%	0,970	-0.12%	0,980	-0.13%	0,978	-0.19%	0,858	-0.80%	0,881	-0.38%	0,877
	(-0.72)		(-0.58)		(-0.53)		(-0.19)		(-0.68)		(-0.32)	
60-80%	0.35%	0,956	0.06%	0,976	-0.28%	0,979	1.22%*	0,912	0.98%	0,924	1.56%**	0,953
	(1.37)		(0.17)		(-0.75)		(1.82)		(1.43)		(2.43)	
bot 80%	0.51%**	0,957	0.73%**	0,972	0.79%***	0,973	-0.54%	0,922	-0.04%	0,931	-0.61%	0,939
	(2.17)		(2.48)		(2.70)		(-0.24)		(-0.04)		(-0.64)	
Long/Short	-0.52%	-0,024	-0.88%*	0,332	-0.84%*	0,332	-1.78%	0,194	-0.89%	0,383	-0.76%	0,393
	(-1.09)		(-1.88)		(-1.99)		(-0.88)		(-0.66)		(-0.54)	
Governance												
top 20%	-0.18%	0,964	0.15%	0,974	0.01%	0,971	-0.22%	0,877	-0.97%	0,919	0.59%	0,920
	(-0.54)		(0.59)		(0.03)		(-0.15)		(-1.04)		(0.52)	
20-40%	0.24%	0,965	0.33%	0,977	0.32%	0,976	1.54%*	0,937	1.62%*	0,950	1.05%	0,953
	(0.89)		(1.09)		(1.09)		(1.77)		(1.88)		(1.24)	
40-60%	0.20%	0,959	-0.17%	0,978	-0.01%	0,978	0.61%	0,904	-0.21%	0,924	0.50%	0,922
	(0.96)		(-0.82)		(-0.06)		(0.83)		(-0.27)		(0.75)	
60-80%	-0.28%	0,958	-0.55%*	0,970	-0.58%*	0,969	-1.41%	0,929	-1.50%	0,951	-2.37%**	0,946
	(-1.20)		(-1.95)		(-1.82)		(-1.26)		(-1.38)		(-2.59)	
bot 80%	1.07%*	0,939	0.86%*	0,958	0.99%**	0,957	0.37%	0,926	-0.28%	0,956	-0.72%	0,968
	(2.03)		(1.82)		(2.13)		(0.40)		(0.67)		(-1.02)	
Long/Short	-0.94%	0,265	-0.71%	0,550	-0.98%	0,496	-0.60%	0,015	-0.69%	0,480	1.31%	0,529
	(-1.42)		(-1.18)		(-1.47)		(-0.26)		(-0.99)		(0.79)	

Table 8: ESG multi-factor performance statistics (Regression Coefficients)

This table shows the performance statistics of various ESG-sorted portfolios of U.S. Investment Grade (IG) and U.S. High Yield (HY) corporate bonds over the period January 2015 to March 2021. In four column groups, the table shows long-only portfolios that go long the top 20% of bonds, long-only portfolios that go long the bottom 20% of bonds, and long-short portfolios that go long the top 20% and short the bottom 20% of bonds based on Thomson Reuters (TR) ESG Combined scores, Environment scores, Social scores, and Governance scores. The table shows the coefficients from regressions of each portfolio with respect to Model 2, the construction of which is described in Section 6. We annualize the model's Intercept term (alpha). *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence levels, respectively, of two-sided tests on whether the coefficients are different from 0. *t*-statistics with Newey-West standard errors).

	ESG			Environment			Social			Governance		
	top 20%	bot 20%	Long/ Short	top 20%	bot 20%	Long/ Short	top 20%	bot 20%	Long/ Short	top 20%	bot 20%	Long/ Short
	Investment Grade											
Intercept	-0.08%	0.76%***	-0.85%**	-0.26%	0.67%***	-0.93%**	-0.15%	0.73%**	-0.88%*	0.15%	0.86%*	-0.71%
	(-0.38)	(2.92)	(-2.62)	(-0.95)	(3.19)	(-2.34)	(-0.61)	(2.48)	(-1.88)	(0.59)	(1.82)	(-1.18)
Equity Market	0.00	0.00	-0.00	-0.01	0.01*	-0.01	-0.00	0.00	-0.00	0.00	-0.00	0.00
	(0.04)	(0.64)	(-0.48)	(-1.04)	(1.71)	(-1.57)	(-0.08)	(0.34)	(-0.26)	(0.13)	(-0.44)	(0.36)
Low Risk (c)	0.07	0.03	0.03	0.04	0.06	-0.02	0.19**	-0.10*	0.29***	-0.19***	0.09	-0.28***
	(0.84)	(0.64)	(0.42)	(0.72)	(1.23)	(-0.25)	(2.60)	(-1.88)	(3.25)	(-3.59)	(1.37)	(-2.97)
Market	1.14***	0.90***	0.24***	1.13***	0.84***	0.29***	1.02***	0.95***	0.07**	1.18***	0.82***	0.36***
	(69.86)	(50.45)	(11.05)	(27.87)	(43.38)	(5.31)	(49.18)	(36.11)	(2.58)	(37.12)	(17.48)	(5.72)
Momentum (6)	0.04***	0.07***	-0.03*	0.01	0.05***	-0.04*	0.01	0.02*	-0.02	-0.01	0.04*	-0.05
	(3.38)	(6.00)	(-1.71)	(0.35)	(4.10)	(-1.71)	(0.63)	(1.88)	(-0.85)	(-0.41)	(1.93)	(-1.42)
Size	-0.20***	0.09**	-0.30***	-0.22***	0.13***	-0.36***	-0.12***	0.12**	-0.24***	-0.17***	0.07	-0.24***
	(-5.41)	(2.51)	(-7.13)	(-6.84)	(4.23)	(-6.26)	(-3.86)	(2.30)	(-3.26)	(-4.38)	(1.49)	(-3.62)
TERM	-0.03	-0.00	-0.02	-0.03	0.01	-0.05	-0.01	-0.04	0.03	-0.03	-0.01	-0.02
	(-1.17)	(-0.24)	(-0.71)	(-1.66)	(0.60)	(-1.25)	(-0.56)	(-1.41)	(0.83)	(-1.12)	(-0.20)	(-0.45)
Value	-0.01	0.03	-0.05***	-0.01	-0.00	-0.00	-0.01	0.01	-0.02	-0.04	0.05*	-0.09**
	(-0.76)	(1.63)	(-2.69)	(-0.21)	(-0.21)	(-0.06)	(-0.74)	(0.65)	(-1.02)	(-1.46)	(1.92)	(-2.44)
adj R^2	0.984	0.979	0.583	0.982	0.978	0.624	0.980	0.972	0.332	0.974	0.958	0.550
	High Yield											
Intercept	-2.31%**	0.91%	-3.22%**	-3.16%***	0.82%	-3.99%***	-0.93%	-0.04%	-0.89%	-0.97%	0.58%	-1.55%
	(-2.48)	(1.27)	(-2.24)	(-3.32)	(0.87)	(-2.70)	(-1.49)	(-0.04)	(-0.66)	(-1.04)	(0.67)	(-0.99)
Equity Market	0.04**	-0.01	0.06**	0.03	-0.02	0.05*	0.03**	0.00	0.02	0.00	-0.00	0.00
	(2.46)	(-1.10)	(2.28)	(1.65)	(-1.18)	(1.78)	(2.00)	(0.20)	(0.97)	(0.12)	(-0.08)	(0.13)
Low Risk (c)	0.35***	0.07	0.27***	0.31***	0.04	0.27***	0.31***	0.02	0.29***	-0.01	0.17***	-0.18
	(6.09)	(1.15)	(2.83)	(5.53)	(0.57)	(2.69)	(4.57)	(0.27)	(2.85)	(-0.06)	(2.92)	(-1.57)
Market	1.05***	0.88***	0.17***	1.14***	0.96***	0.19***	1.15***	0.94***	0.21***	0.82***	0.94***	-0.12
	(26.45)	(23.76)	(3.20)	(26.74)	(21.19)	(4.28)	(29.21)	(23.24)	(4.25)	(13.32)	(22.22)	(-1.32)
Momentum (6)	0.03**	0.01	0.03	0.04**	0.03	0.00	0.02	0.02	0.00	-0.02	0.02	-0.04
	(2.04)	(0.43)	(1.12)	(2.25)	(1.64)	(0.17)	(1.51)	(0.99)	(0.10)	(-0.59)	(1.20)	(-0.87)
Size	-0.01	-0.04	0.04	-0.01	-0.03	0.02	0.06	-0.11***	0.17***	-0.00	-0.04	0.03
	(-0.26)	(-1.63)	(0.84)	(-0.23)	(-1.00)	(0.45)	(1.58)	(-2.93)	(3.73)	(-0.04)	(-1.30)	(0.64)
TERM	0.06	-0.03	0.09	0.03	-0.00	0.03	0.05	-0.05	0.10	-0.03	-0.01	-0.03
	(0.82)	(-0.69)	(0.93)	(0.44)	(-0.00)	(0.28)	(0.91)	(-0.66)	(1.25)	(-0.41)	(-0.16)	(-0.24)
Value	0.09**	-0.08***	0.17***	0.08*	-0.10***	0.18***	0.03	-0.06*	0.09	0.25***	-0.13***	0.38***
	(2.51)	(-4.14)	(3.81)	(1.93)	(-3.67)	(3.55)	(0.72)	(-1.70)	(1.45)	(5.01)	(-4.28)	(6.58)
adj R^2	0.948	0.961	0.351	0.946	0.948	0.403	0.962	0.931	0.383	0.919	0.956	0.480

The performance of Low-Risk may also help to explain the relative underperformance of top HY quintile portfolios. As Houweling and van Zundert (2017) and Slimane et al. (2018) suggest, Low-Risk is a clear “bad times” factor, and as such performs best relative to other factor portfolios during crisis periods. According to Table 7, our Low-Risk portfolio does not exhibit the strong performance that it does in Houweling and van Zundert (2017) (see Table 7). We suggest that this could be a product of our sample period, which spans the latter half of a massive bull market in fixed income, and which contains only a few months of market turmoil.[105] We find that all of our top quintile portfolios display large and positive coefficients on Low-Risk. As we explain in greater detail in section 6.3d, the top Governance portfolio, which has the largest and most statistically significant coefficient on Low-Risk, performs best during the market crash in 2020. This is consistent with Amiraslani et al. (2019), who find the benefits (costs) of strong (weak) CSR to be contained within crisis periods. In a sample with proportionately more months of market weakness, it is possible that we would see an improvement in the average financial performance of our top quintile portfolios in HY.

6.3.3 Initial robustness tests

For studies that analyze the performance of SRI/ESG funds (e.g. Henke, 2016), multifactor models are the most effective (and in some cases only) tools with which researchers can account for variables other than ESG that may influence returns. That we construct synthetic portfolios, however, provides us the flexibility to conduct tests that are even more robust. We structure these tests around those in Fama and French (2008), Houweling and van Zundert (2017) and Dorfleitner et al. (2020). These include a series of bivariate sorts, in which we first group bonds into five quintiles based on size, credit rating or maturity, and then take bonds from each quintile with the 20% highest/lowest ESG scores. We also split our bonds into two groups based on the aforementioned variables, and form two sets of five quintile portfolios with respect to each ESG rating. We repeat our analyses from the previous sections on these new portfolios, and find that, despite the distinct trends we see in Table 6, controlling for various bond characteristics does not alter any of our conclusions so far. [106] In Table 9, we report performance statistics of long-short portfolios constructed using the first double-sorting procedure.[107]

That our results are robust to these tests further distinguishes them from those in the existing literature. For example, Oikonomou et al. (2014) find the spread-tightening effect of CSR to be stronger at longer maturities. Similarly, Hoepner and Nilsson (2020) show that the negative relationship between bond returns and both CSR strengths and concerns diminishes in portfolios with long average durations. Among equally-weighted ESG equity portfolios, Dorfleitner et al. (2020) find that those with the best performance not only have much lower average ratings, but are comprised of stocks from smaller companies. After they control for differences in market capitalization, they find that the abnormal returns to low-rated portfolios disappear. The lack of any meaningful differences between the returns of our base quintile portfolios and those we form with the 5x2 and 2x5 sorting procedures could be (at least in part) a byproduct of our representative bond filter. Oikonomou et al.

[105] Most notably during the end of 2018 and during the Coronavirus-induced market crash of 2020. While the latter was remarkable in regards to both its speed and severity, it only lasted a few months.

[106] We do see some minor differences. After controlling for years to maturity, for example, the bottom HY Governance portfolio produces a negative alpha almost twice that of the base portfolio. While we cannot fit results from all portfolios into a reasonably sized table, we can confirm that the outperformances of the portfolios of the 20%-40% rated bonds in HY also remains unchanged.

[107] We occasionally reference results from the second sorting procedure, which plays a much larger role in section 6.4a.

(2014) show that the influence of maturity on the returns to ESG-sorted portfolios is only apparent in bonds with maturities greater than twenty years. Since our filter creates a sample that is biased towards bonds with intermediate durations, it is possible that we miss these maturity effects altogether. Similarly, our division of IG and HY bonds reduces the variation in credit ratings. For some researchers, this division is the extent to which they control for differences in credit quality (see the robustness tests featured in Bai et al., 2018). In any case, we suggest that these tests go a long way towards demonstrating the robustness of our results to changes in prominent bond characteristics. More importantly, they support our assertion that, in terms of the performance of ESG-sorted bond portfolios, the most relevant distinction is that between IG and HY market segments.

6.3.4 Returns to ESG-sorted portfolios across time

The two main theories regarding the time dependency of the returns to ESG-sorted portfolios, the errors-in-expectations and the shunned-stock hypotheses, make two very straightforward and testable predictions. According to errors-in-expectations, any abnormal returns attributable to ESG factors dissipate over time as investors learn to accurately incorporate ESG into security analysis (Derwall et al., 2011). The shunned-stock effect, on the other hand, asserts that investors' non-pecuniary preferences for securities with high ESG ratings creates a shortage of demand for poorly rated securities, driving up their expected returns over time. Derwall et al. (2011) assert that, while the two effects oppose one another other in the short term, only the shunned-stock effect persists in the long run.[108] Consistent with greater learning and attention to ESG, subsequent researchers show that once profitable ESG strategies no longer generate abnormal returns (Cremars, 2007, Halbritter and Dorfleitner, 2015, Pereira, 2019, Dorfleitner et al., 2020). Pereira (2019) provides the first and only examination of how well these theories explain the returns to corporate bonds, and finds evidence of both shunned-stock and errors-in-expectations in the returns to ESG-sorted portfolios of European bonds.

We predict that we will see evidence of the shunned-stock effect in IG, which would help to reconcile our results (the overperformance of bottom quintile portfolios) with those in the existing literature (which finds that high ESG bonds outperformed low ESG bonds in the years prior to the present sample period).[109] Since our results in HY appear to align more closely with risk-based explanations, we anticipate time-series trends in the returns of HY portfolios that conform to the errors-in-expectations hypothesis. We follow Pereira et al. (2019) in using both expanding and rolling window analyses to examine the evolution of the returns to our ESG-sorted portfolios over time. Although these methods allow us to better identify any potential time-series trends compared to a split sample approach, our relatively narrow sample period will nevertheless restrict the strength of any inferences we make. Tables 10 and 11 display the alphas, adjusted t-statistics and adjusted R-squared values from our multivariate regressions with expanding and rolling windows respectively. To conserve space, we only include the results for the top, bottom and long-short portfolios formed using our standard 20% cut-off. As in earlier sections, our results are mostly unchanged when we implement a 10% cut-off rate instead. The intercept statistics are from regressions on Model 2, which we find to be the most robust in terms of capturing the time-series variation in the returns to our quintile portfolios.[110]

[108] They attribute this to the market's gradual elimination of mispricings as opposed to the stickiness of investor preferences.
[109] Derwall and Bauer (2009), Henke (2016) and Pereira (2019).
[110] To credit the work of Bai et al. (2018), we do find that the bond-specific factors from Model 3 perform slightly better in the cross-section than those from Model 2. Our two-pass cross-sectional regressions are rather rough, however, as the LinearModels package in Python does not implement the Shanken correction.

Table 9: Performance of ESG-sorted long-short portfolios with controls on Size, Credit Ratings and Maturity

This table shows the performance statistics of ESG-sorted portfolios for U.S. Investment Grade and U.S. High Yield corporate bonds over the period January 2015 to March 2021. We construct base-case long-short portfolios that go long the top 20% and short the bottom 20% of bonds with respect to their issuer's Thomson Reuters (TR) ESG score, Environment score, Social score, and Governance score. We construct size-, credit- and duration-neutral portfolios by first creating long-short portfolios per quintile of debt outstanding, rating group (AAA/AA+/AA/AA-/A+, A/A-, BBB+, BBB, BBB-, BB+/BB, BB-/B+/B), or quintile of [years to maturity / duration] and then combining all groups to form the final portfolio. The table reports the annualized alpha of each portfolio with respect to two multi-factor models, Models 1 and 2, described in Section 6.2, as well as the models' adjusted R^2. *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence levels, respectively, of two-sided tests on whether the alphas are different from 0. *t*-statistics with Newey-West standard errors. We measure corporate bond returns as excess returns versus maturity-matched Treasuries.

	Investment Grade				*High Yield*			
	Model 1		**Model 2**		**Model 1**		**Model 2**	
	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2
ESG								
base	-0.78%**	0.222	-0.85%**	0.583	-2.18%	0.208	-3.22%**	0.351
	(-2.30)		(-2.62)		(-1.42)		(-2.24)	
size neutral	-0.57%*	0.403	-0.79%***	0.534	-1.52%	0.358	-2.19%*	0.434
	(-1.83)		(-2.80)		(-1.34)		(-1.88)	
credit neutral	-0.38%	0.304	-0.36%	0.564	-1.22%	0.048	-2.33%**	0.273
	(-1.32)		(-1.18)		(-1.16)		(-2.38)	
duration neutral	-0.76%*	0.006	-0.79%*	0.425	-1.53%	0.021	-2.68%***	0.272
	(-1.68)		(-1.88)		(-1.52)		(-2.85)	
Environment								
base	-1.04%***	0.314	-0.93%**	0.624	-3.18%*	0.311	-3.99%***	0.403
	(-2.67)		(-2.34)		(-1.96)		(-2.70)	
size neutral	-0.69%**	0.366	-0.90%**	0.481	-2.65%**	0.184	-3.59%***	0.326
	(-2.41)		(-2.14)		(-2.21)		(-3.28)	
credit neutral	-0.87%**	0.416	-0.77%*	0.639	-2.19%**	0.239	-3.35%***	0.396
	(-2.48)		(-1.97)		(-2.16)		(-3.42)	
duration neutral	-1.25%***	0.078	-1.34%***	0.566	-2.75%*	0.263	-3.35%***	0.405
	(-3.07)		(-3.29)		(-1.95)		(-2.74)	
Social								
base	-0.52%	-0.024	-0.88%*	0.332	-0.01%	0.194	-0.89%	0.383
	(-1.09)		(-1.88)		(-0.01)		(-0.66)	
size neutral	-0.58%	0.008	-0.93%**	0.106	-0.52%	0.347	-1.45%	0.550
	(-1.47)		(-2.55)		(-0.49)		(-1.27)	
credit neutral	-0.38%	-0.030	-0.57%*	0.097	-0.32%	0.339	-1.05%	0.418
	(-1.09)		(-1.84)		(-0.31)		(-0.93)	
duration neutral	-0.67%	0.067	-0.82%*	0.347	0.33%	0.087	-0.51%	0.300
	(-1.44)		(-1.88)		(0.30)		(-0.41)	
Government								
base	-0.94%	0.365	-0.71%	0.550	-0.60%	0.015	-1.55%	0.480
	(-1.42)		(-1.18)		(-0.26)		(-0.99)	
size neutral	-0.66%	0.433	-0.56%	0.509	-2.29%	0.109	-1.83%**	0.393
	(-1.03)		(-0.91)		(-1.53)		(-2.62)	
credit neutral	-0.94%	0.455	-0.63%	0.570	-0.71%	0.016	-1.57%	0.428
	(-1.28)		(-0.98)		(-0.38)		(-1.12)	
duration neutral	-1.14%**	0.228	-1.12%**	0.446	-0.89%	-0.036	-1.32%	0.166
	(-2.27)		(-2.12)		(-0.56)		(-1.02)	

As we predicted, the performance of the bottom Combined quintile in IG is consistent with the shunned-stock effect (Derwall et al. 2011). This is especially apparent in Table 11, where from each two-year interval to the next, the bottom portfolio's alpha rises in economic and statistical significance. Bottom quintiles sorted on individual pillar scores mostly conform to this trend as well, but each portfolio breaks the pattern once (the most egregious case of which is the slightly negative alpha on the bottom Governance portfolio during the 2018 – 2020 period). Ultimately, all of the bottom portfolios in IG exhibit the strongest, positive alphas in the 2019 – 2021 period, some of which would be quite spectacular if they were based on more than just two years of data (e.g. the bottom Governance portfolio produces an alpha of 3.47% with a t-statistic of 6.2). Our results align with those of Pereira (2019), whose analogous bottom portfolios perform best at the end of their sample period. Although we do not find any clear indications of errors-in-expectations, we suggest that this aligns with Derwall's (2011) prediction that errors-in-expectations culminates when top-rated securities no longer outperform the market. If we rely on evidence from Henke (2016) and Pereira (2019) that bonds from issuers with high ESG scores posted substantial, positive alphas prior to the beginning of our sample, our results broadly support the claim that bondholders more effectively discount the ramifications of strong ESG practices (at least insofar as such practices are reflected in the TR score).[111]

Even if we could verify that prices of IG bonds now more fairly represent the valuation effects of ESG, the same would not necessarily be true of HY bonds. As Jostova et al. (2013) note, the HY market is more fragmented and contains a larger portion of private firms, both of which slow the transmission of information into prices. The authors suggest that the presence of a momentum effect in HY but not in IG, which is consistent with our findings in Section 6.1, further indicates a divide with respect to this transmission.[112] We find no patterns in Tables 10 or 11 that clearly support either the shunned-stock or errors-in-expectations hypotheses in HY. If anything, we find that bottom quintile portfolios performed better in the past, which is the opposite of what we expect based on the shunned-stock effect as researchers have traditionally described it (Derwall et al., 2011, Pereira, 2019). Insofar as low ESG scores reflect practices that are beneficial to HY issuers, we could construe this trend as an indicator of errors-in-expectations.[113] Overall, however, time series trends in the returns to HY portfolios are not very distinct, and do not align with those of any previous study of which we are aware. There are some peculiarities at the end of our sample period, such as the strong outperformance of the top Governance portfolio, but we discuss our interpretation of these instances in the next section. At the very least, this analysis displays yet another way in which ESG-sorted portfolios of HY bonds exhibit behavior that diverges from those in IG.

[111] Pereira's (2019) portfolios contain only European bonds. Given evidence that the shunned-stock and errors-in-expectations effects play out at different times in different markets (Badia and Cortez, 2020), the applicability of her findings to ours is admittedly dubious.
[112] Admittedly, our finding of a reversal effect in IG does undermine the extent to which the argument of Jostova et al. (2013) applies in our case.
[113] Although researchers typically refer to errors-in-expectations to explain the financial benefits of ESG, it could theoretically explain negative returns as well.

Table 10: ESG performance in expanding sample

This table shows the performance statistics of various ESG-related strategies for U.S. Investment Grade (IG) and U.S. High Yield (HY) corporate bonds over an expanding sample series from January 2015 to March 2021. The series expands by one year between samples. The left hand side of the table shows long-only portfolios that go long the top 20% of bonds, long-only portfolios that go long the bottom 20% of bonds, and long-short portfolios that go long the top 20% and short the bottom 20% of bonds based on issuers' Thomson Reuters (TR) ESG Combined scores, Environment scores, Social scores, and Governance scores. The table shows annualized alpha statistics for each portfolio with respect to a multi-factor model, Model 2, described in Section 6.2. *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence levels, respectively, of two-sided tests on whether the coefficients are different from 0 (*t*-tests with Newey-West standard errors). We measure corporate bond returns as excess returns versus maturity-matched Treasuries.

	[2015,2017]		[2015,2018]		[2015,2019]		[2015,2020]		[2015,2021]	
	alpha	adj R2	alpha	adj R2	alpha	adj R2	alpha	adj R2	alpha	adj R2
					Investment Grade					
ESG										
top 20%	-0.41%	0.994	-0.49%**	0.993	-0.16%	0.988	-0.13%	0.983	-0.14%	0.984
	(-1.32)		(-2.53)		(-0.86)		(-0.82)		(-0.83)	
bot 20%	0.25%	0.988	0.65%***	0.981	0.45%**	0.977	0.43%**	0.977	0.75%**	0.979
	(1.30)		(3.18)		(2.08)		(2.46)		(2.61)	
Long/ Short	-0.66%	0.505	-1.14%***	0.452	-0.61%	0.355	-0.56%*	0.478	-0.90%**	0.585
	(-1.34)		(-3.36)		(-1.67)		(-1.92)		(-2.49)	
Environment										
top 20%	-0.43%**	0.993	-0.52%***	0.991	-0.19%	0.988	-0.08%	0.981	-0.28%	0.982
	(-2.20)		(-2.84)		(-0.98)		(-0.46)		(-1.12)	
bot 20%	0.49%***	0.990	0.76%***	0.987	0.52%***	0.982	0.41%**	0.976	0.64%**	0.978
	(4.98)		(4.62)		(2.79)		(2.24)		(2.56)	
Long/ Short	-0.92%***	0.621	-1.27%***	0.549	-0.71%*	0.449	-0.49%	0.380	-0.93%**	0.621
	(-4.02)		(-4.07)		(-1.93)		(-1.52)		(-2.07)	
Social										
top 20%	-0.58%*	0.993	-0.38%	0.987	-0.07%	0.983	-0.09%	0.978	-0.22%	0.981
	(-1.85)		(-1.20)		(-0.28)		(-0.39)		(-0.94)	
bot 20%	0.13%	0.972	0.27%	0.971	0.28%	0.968	0.50%**	0.969	0.72%**	0.974
	(0.45)		(0.78)		(1.12)		(2.40)		(2.56)	
Long/ Short	-0.71%	0.209	-0.65%	0.224	-0.35%	0.208	-0.59%	0.191	-0.94%**	0.345
	(-1.29)		(-1.12)		(-0.81)		(-1.63)		(-2.07)	
Governance										
top 20%	0.02%	0.989	-0.03%	0.988	0.12%	0.987	0.27%*	0.979	0.04%	0.981
	(0.18)		(-0.22)		(0.88)		(1.68)		(0.19)	
bot 20%	0.42%	0.990	0.78%**	0.986	0.78%***	0.982	0.55%	0.964	0.90%*	0.960
	(1.23)		(2.12)		(2.78)		(1.48)		(1.72)	
Long/ Short	-0.41%	0.577	-0.81%**	0.457	-0.66%*	0.436	-0.28%	0.527	-0.86%	0.608
	(-1.36)		(-2.39)		(-1.84)		(-0.62)		(-1.24)	
					High Yield					
ESG										
top 20%	-3.08%*	0.869	-3.88%**	0.846	-2.14%*	0.829	-2.23%*	0.823	-2.28%**	0.948
	(-1.78)		(-2.39)		(-1.86)		(-1.94)		(-2.14)	
bot 20%	3.21%**	0.824	2.25%*	0.833	1.25%	0.842	1.51%*	0.852	1.04%	0.963
	(2.35)		(1.85)		(1.31)		(1.75)		(1.39)	
Long/ Short	-6.29%**	0.290	-6.13%**	0.317	-3.39%*	0.239	-3.74%*	0.243	-3.32%*	0.351
	(-2.24)		(-2.47)		(-1.78)		(-2.00)		(-1.94)	
Environment										
top 20%	-2.87%	0.848	-3.65%**	0.822	-2.48%**	0.819	-2.66%**	0.821	-3.13%***	0.946
	(-1.60)		(-2.30)		(-2.31)		(-2.49)		(-2.94)	
bot 20%	1.82%	0.771	1.60%	0.788	0.77%	0.807	1.20%	0.823	1.02%	0.955
	(0.95)		(1.22)		(0.78)		(1.31)		(1.05)	
Long/ Short	-4.69%	0.399	-5.25%**	0.380	-3.25%*	0.321	-3.86%**	0.315	-4.15%**	0.406
	(-1.37)		(-2.11)		(-1.82)		(-2.19)		(-2.30)	
Social										
top 20%	0.12%	0.907	-0.80%	0.872	-0.45%	0.864	-0.41%	0.854	-0.88%**	0.961
	(0.09)		(-0.68)		(-0.60)		(-0.67)		(-2.29)	
bot 20%	-0.59%	0.700	0.00%	0.741	-0.03%	0.768	0.44%	0.797	0.03%	0.931
	(-0.32)		(0.00)		(-0.03)		(0.47)		(0.03)	
Long/ Short	0.71%	0.051	-0.81%	0.089	-0.42%	0.152	-0.85%	0.235	-0.91%	0.381
	(0.23)		(-0.33)		(-0.32)		(-0.65)		(-0.76)	
Governance										
top 20%	-2.87%**	0.935	-3.33%***	0.917	-1.24%	0.870	-1.94%**	0.861	-1.05%	0.920
	(-2.15)		(-4.22)		(-1.36)		(-2.55)		(-1.46)	
bot 20%	3.69%**	0.774	2.59%*	0.768	1.31%	0.774	1.51%	0.801	0.65%	0.957
	(2.66)		(2.00)		(1.12)		(1.52)		(0.75)	
Long/ Short	-6.56%**	0.749	-5.92%***	0.716	-2.55%	0.578	-3.45%**	0.525	-1.70%	0.486
	(-2.51)		(-3.15)		(-1.34)		(-2.11)		(-1.14)	

Table 11a: ESG performance in rolling sample (Investment Grade)

This table shows the performance statistics of various ESG-related strategies for U.S. Investment Grade (IG) corporate bonds over a rolling sample series from January 2015 to March 2021. Each interval in the series spans a period of two years with the exception of the last period (COVID-19) which spans January 2020 to March 2021. The series increments by one year between intervals. The left hand side of the table shows long-only portfolios that go long the top 20% of bonds, long-only portfolios that go long the bottom 20% of bonds, and long-short portfolios that go long the top 20% and short the bottom 20% of bonds based on Thomson Reuters (TR) ESG Combined scores, Environment scores, Social scores, and Governance scores. The table shows annualized alpha statistics for each portfolio with respect to a multi-factor model, Model 2, described in Section 6.2. *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence levels, respectively, of two-sided tests on whether the coefficients are different from 0 (*t*-tests with Newey-West standard errors). We measure corporate bond returns as excess returns versus maturity-matched Treasuries.

	[2015,2017)		**[2016,2018)**		**[2017,2019)**		**[2018,2020)**		**[2019,2021)**		**COVID-19**	
	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2
	Investment Grade											
ESG												
top 20%	-0.41%	0.994	-0.78%***	0.994	-0.09%	0.974	0.19%	0.976	-0.80%**	0.984	-1.38%	0.991
	(-1.32)		(-4.27)		(-0.32)		(1.02)		(-2.75)		(-1.63)	
bot 20%	0.25%	0.988	0.21%	0.982	0.22%*	0.965	0.30%	0.973	1.47%***	0.979	4.39%***	0.983
	(1.30)		(0.71)		(1.87)		(1.47)		(4.02)		(16.36)	
Long/ Short	-0.66%	0.505	-0.99%**	0.705	-0.31%	0.594	-0.11%	0.700	-2.27%***	0.618	-5.77%***	0.734
	(-1.34)		(-2.37)		(-0.83)		(-0.56)		(-3.93)		(-5.70)	
Environ-ment												
top 20%	-0.43%**	0.993	-0.61%***	0.989	-0.16%	0.982	0.04%	0.981	-1.36%**	0.977	-2.30%	0.988
	(-2.20)		(-5.86)		(-0.71)		(0.17)		(-2.52)		(-0.99)	
bot 20%	0.49%***	0.990	0.31%	0.985	0.42%***	0.964	0.21%	0.972	1.82%***	0.967	3.81%***	0.981
	(4.98)		(1.20)		(4.55)		(1.50)		(2.96)		(8.86)	
Long/ Short	-0.92%***	0.621	-0.93%***	0.658	-0.58%**	0.560	-0.16%	0.626	-3.18%**	0.643	-6.11%*	0.852
	(-4.02)		(-3.37)		(-2.37)		(-0.66)		(-2.92)		(-2.25)	
Social												
top 20%	-0.58%*	0.993	-0.50%	0.979	-0.05%	0.965	-0.12%	0.979	-0.54%	0.972	0.41%	0.980
	(-1.85)		(-1.30)		(-0.32)		(-0.83)		(-1.55)		(0.69)	
bot 20%	0.13%	0.972	-0.61%	0.967	0.10%	0.951	0.63%*	0.967	2.12%***	0.985	3.59%***	0.972
	(0.45)		(-1.05)		(0.35)		(2.12)		(6.86)		(5.99)	
Long/ Short	-0.71%	0.209	0.11%	0.359	-0.15%	0.338	-0.75%**	0.389	-2.65%***	0.457	-3.18%***	0.529
	(-1.29)		(0.12)		(-0.35)		(-2.60)		(-4.59)		(-8.50)	
Governance												
top 20%	0.02%	0.989	-0.23%*	0.994	0.22%*	0.978	0.17%	0.975	-1.25%**	0.977	-1.68%	0.962
	(0.18)		(-2.08)		(1.84)		(0.66)		(-2.72)		(-0.99)	
bot 20%	0.42%	0.990	0.41%	0.986	0.61%***	0.956	-0.24%	0.954	3.47%***	0.985	3.27%***	0.988
	(1.23)		(1.74)		(4.61)		(-1.26)		(6.18)		(6.54)	
Long/ Short	-0.41%	0.577	-0.64%**	0.540	-0.39%*	0.463	0.41%*	0.707	-4.72%***	0.830	-4.95%*	0.705
	(-1.36)		(-2.21)		(-1.80)		(1.97)		(-4.86)		(-2.30)	

Table 11b: ESG performance in rolling sample (High Yield)

This table shows the performance statistics of various ESG-related strategies for U.S. High Yield (HY) corporate bonds over a rolling sample series from January 2015 to March 2021. Each interval in the series spans a period of two years with the exception of the last period (COVID-19) which spans January 2020 to March 2021. The series increments by one year between intervals. The left hand side of the table shows long-only portfolios that go long the top 20% of bonds, long-only portfolios that go long the bottom 20% of bonds, and long-short portfolios that go long the top 20% and short the bottom 20% of bonds based on Thomson Reuters (TR) ESG Combined scores, Environment scores, Social scores, and Governance scores. The table shows annualized alpha statistics for each portfolio with respect to a multi-factor model, Model 2, described in Section X. *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence levels, respectively, of two-sided tests on whether the coefficients are different from 0. t-statistics with Newey-West standard errors. We measure corporate bond returns as excess returns versus maturity-matched Treasuries.

	[2015,2017]		**[2016,2018]**		**[2017,2019]**		**[2018,2020]**		**[2019,2021]**		**COVID-19**	
	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2
							High Yield					
ESG												
top 20%	-3.08%*	0.869	0.00%	0.904	-0.99%	0.755	-1.28%**	0.827	-0.60%	0.978	0.51%	0.982
	(-1.78)		(0.00)		(-1.39)		(-2.32)		(-0.46)		(0.54)	
bot 20%	3.21%**	0.824	0.05%	0.884	0.18%	0.920	0.34%	0.880	2.02%**	0.983	-0.20%	0.989
	(2.35)		(0.07)		(0.43)		(0.46)		(2.36)		(-0.23)	
Long/ Short	-6.29%**	0.290	-0.05%	0.383	-1.17%*	0.455	-1.62%	0.241	-2.61%	0.251	0.71%	0.309
	(-2.24)		(-0.03)		(-1.91)		(-1.30)		(-1.33)		(0.60)	
Environment												
top 20%	-2.87%	0.848	0.16%	0.875	-1.82%**	0.784	-2.08%***	0.888	-1.78%	0.983	0.13%	0.981
	(-1.60)		(0.16)		(-2.65)		(-5.28)		(-1.70)		(0.12)	
bot 20%	1.82%	0.771	0.02%	0.805	0.06%	0.902	0.45%	0.920	2.66%**	0.979	-0.19%	0.970
	(0.95)		(0.02)		(0.14)		(1.06)		(2.57)		(-0.15)	
Long/ Short	-4.69%	0.399	0.14%	0.312	-1.88%***	0.477	-2.54%***	0.445	-4.44%***	0.449	0.32%	0.424
	(-1.37)		(0.07)		(-4.53)		(-4.30)		(-3.58)		(0.24)	
Social												
top 20%	0.12%	0.907	1.04%	0.885	-1.51%**	0.772	-1.16%***	0.844	-1.79%*	0.981	-0.18%	0.986
	(0.09)		(1.01)		(-2.78)		(-4.23)		(-1.92)		(-0.19)	
bot 20%	-0.59%	0.700	-2.34%*	0.825	0.08%	0.910	0.67%	0.871	1.11%	0.963	-1.07%	0.958
	(-0.32)		(-1.77)		(0.14)		(0.54)		(0.89)		(-0.44)	
Long/ Short	0.71%	0.051	3.38%	0.152	-1.59%***	0.583	-1.84%	0.392	-2.90%	0.425	0.89%	0.216
	(0.23)		(1.55)		(-5.22)		(-1.39)		(-1.44)		(0.33)	
Governance												
top 20%	-2.87%**	0.935	-1.13%	0.933	-0.05%	0.810	-0.70%	0.874	2.69%***	0.987	2.72%**	0.981
	(-2.15)		(-1.47)		(-0.14)		(-1.69)		(4.65)		(2.48)	
bot 20%	3.69%**	0.774	0.69%	0.805	-0.52%	0.863	-0.51%	0.903	0.60%	0.991	0.68%	0.993
	(2.66)		(1.13)		(-0.93)		(-1.01)		(0.86)		(1.47)	
Long/ Short	-6.56%**	0.749	-1.81%	0.789	0.47%	0.245	-0.19%	-0.026	2.09%**	0.778	2.04%	0.651
	(-2.51)		(-1.49)		(1.25)		(-0.26)		(2.54)		(1.64)	

6.3.5 Crisis-period performance

One of the most potent bull markets in decades for corporate bonds dominates our sample. That we use maturity-matched excess returns dampens some of the tremendous duration gains bonds experienced during this period. Nevertheless, the average Sharpe ratios of our IG (0.71) and HY (0.63) market portfolios are slightly more than double those of Houweling and van Zundert (2017) and Bai et al. (2018) (whose samples run from 2000/2001 to 2015).[114] While the crash in financial markets following the outbreak of SARS-CoV-2 provides us with a period of extreme market turmoil, this rout was short-lived, and credit spreads took only seven months to return to pre-crisis levels (a feat that took GFC-era spreads three years).[115] To hone in on the performance of our ESG-sorted portfolios during the crisis, we append regression results from a one-year period (February 2020 to February 2021) to the results from our two-year rolling windows in Table 11. Given some of the methodological concerns related to an examination of such a short time period, however, we do not place much credence in the precision of our model's estimates. Instead, we look for substantial deviations in the performances of our portfolios during crisis and non-crisis periods.

Following the GFC, researchers largely agreed that high ESG securities experienced various advantages during the market panic (Cornett et al., 2016, Lins et al., 2017). These included a stable investor base (El Ghoul et al, 2017, Demers et al., 2020), more supportive stakeholders (Sacconi and Degli Antoni 2011, Lins et al., 2017), and decreased risks of agency problems and bankruptcy (Amiraslani et al., 2019, Albuquerque et al., 2020). Scholars have long examined the insurance-like effects of social capital (Godfrey et al., 2009), which firms appear to accumulate over years of positive interactions with society broadly (legitimacy theory), and with stakeholders in particular (stakeholder theory) (Menz, 2010). Moreover, researchers increasingly emphasize ESG's role in downside risk and volatility reduction (Grossner, 2017, Hoepner et al., 2019). Albuquerque et al. (2020) establish a theoretical framework in which, through improved product differentiation, a firm's investments in CSR reduce its exposure to systematic risks. A number of papers that compare the performance of ESG-integrated mutual funds to that of the S&P 500 also stress that a firm's sustainability record appears to influence its sensitivity to market swings (Solar-Dominguez and Matallin-Saez, 2016, Ouchen, 2021), and that its risk mitigation effects are even stronger during recessionary periods (Henke, 2016, Leite and Cortez, 2016). Regarding the impact of ESG during the COVID-19-induced crash, however, empirical studies have yielded mixed results. While Pastor and Vorsatz (2020) find that flows to SRI funds were less volatile than those to conventional funds, Döttling and Kim (2020) show that, after controlling for fund characteristics, SRI funds actually experienced sharper declines in retail flows during the pandemic. In possibly the most comprehensive examination of the effects of ESG during both the GFC and the COVID-19 crash, Demers et al. (2020) show that, after controlling for a wide array of market and accounting-based indicators, ESG played no discernible role in equity performance.[116] Lastly, to the extent that ESG investments reflect an inefficient use of scarce internal resources, Dorfleitner et al. (2020) note that spending on ESG during a financial crisis can prove detrimental to a firm's financial stability (e.g. the "overinvestment view", Barber, 2007, Barth et al., 2019).

---

[114] Sharpe ratios fall to 0.66 and 0.55 if we consider the performance of all bonds in our dataset and not just those that remain after we implement the representative/liquid bond filter and delete all bonds without ESG scores.

[115] https://fred.stlouisfed.org/series/BAA10Y

[116] They do find, however, that value creation through intangible asset formation appears to have played a positive role.

It is within these varied explanations that we make sense of our results. As we expect based on our analysis in the previous section, our IG and HY portfolios behave differently during the crisis. In IG, the overperformance of our bottom portfolios relative to our top portfolios reaches a peak during the 2019 – 2021 and COVID-19 periods (see Table 11). So much so that the alphas of long-short portfolios formed on the Combined, Environmental, Social and Governance scores rise to statistically significant values of -2.3% (-5.8%), -3.2% (-6.1%), -2.65% (-3.2%) and -4.7% (-5%) during the 2019 – 2021 (COVID-19) periods respectively.[117] As we discussed previously, this is due in large part to the phenomenal returns of the bottom portfolios at the end of our sample period. We do not have advanced methods of isolating the potential influence of the shunned-stock effect from that of the COVID-19 crisis, but we can test whether the alphas of our quintile portfolios during 2019 differ from those of 2020. Based on Model 2, the alphas of the top portfolios are statistically indistinguishable from one another, but those of the bottom Environmental and Governance portfolios are materially higher during the COVID-19 period.[118] If we disregard the bottom portfolios, however, our results appear to align with those of Demers et al. (2020), and, although our crisis periods differ, those of Barth et al. (2019) and Halling (2020), both of whom fail to detect meaningful differences in the relationship between ESG and bond/CDS spreads during the GFC.

We observe the opposite trend in HY, such that the outperformance of bottom quintile portfolios relative to top portfolios diminishes in the 2019 – 2021 and COVID-19 periods. In fact, in terms of average excess returns, each of the top portfolios outperform their lower-rated peers during the height of the COVID-19 crisis from February 2021 to August 2021. The top quintile portfolio based on the Combined score posts cumulative excess returns that are nearly double those of the bottom portfolio (12.4% vs. 6.4%). As we see in Table 8, however, top quintile portfolios exhibit higher market betas, and once we account for this and other risk factors in Model 2, the top portfolios' relative outperformance becomes much less pronounced. The top Governance portfolio is a notable exception and, even with Model 2, produces statistically significant alphas of 2.7% and 3.4% across the 2019 – 2021 and COVID-19 periods respectively. Table 11 further illustrates the period-specific nature of the returns to our Governance portfolios, with the top quintile portfolio posting an alpha of -1.83% in a sample that excludes 2020, but one that is indistinguishable from zero when we consider the full sample.

We suggest that these results are quite remarkable in the context of our prediction that agency risks drive the patterns we observe in the returns to HY bonds. They align very closely with those of Amiraslani et al. (2019), who show that CSR's role in reducing credit spreads during the GFC was related to a firm's exposure to the agency costs of equity and debt. Specifically, the authors find that default-prone firms with few tangible assets almost wholly accounted for the spread-tightening effect associated with strong CSR practices during the crisis. The authors do not examine the role of individual ESG pillars, but the qualities of CSR to which the authors attribute this effect mostly match definitions of corporate governance. In the equities literature, numerous researchers have demonstrated that strong governance improves shareholder value by reducing agency problems and improving the alignment of manager and shareholder objectives (Gerard, 2018). However, as Cremars (2007)

[117] We also look at average excess returns of the bottom portfolios compared to the top portfolios from February 2020 to November 2020, a period during which volatility, as measured by the VIX, was at its highest. The bottom portfolio actually underperforms on an absolute basis (-1.2% compared to the top portfolio), but displays considerably lower volatility (weekly standard deviation of 2.1% compared to 2.9% of the top portfolio). This is consistent with our results in Table 8, in which the bottom portfolios exhibit lower loadings on the market portfolio.

[118] According to regressions on Model 1, the top Social portfolio appears to out-perform during the crisis, but we also find that the explanatory power of Model 1 falls even more during this period.

notes, this high degree of alignment, which often leads to greater shareholder control, can have negative ramifications for bondholders (i.e. cash diversion, leveraged buyouts). The author shows that, while many governance mechanisms are associated with lower bond yields (board independence, highly dispersed institutional ownership), those that shift the balance of power away from management and towards shareholders push bond yields higher. The TR Governance score reflects the former and not the latter, and includes measures of management strength, stakeholder engagement and executive pay, the latter two of which researchers identify as among the most important factors in reducing expected agency costs (Amiraslani et al., 2019, Barney, 2019, Hoepner and Nilsson, 2020).[119] This is also consistent with Salvi et al. (2019), who attribute the negative relationship between CSR and credit risk to compensation practices, management's alignment with firm objectives and the degree to which ESG is integrated within a firm's overall strategy. Therefore, in addition to further distinguishing the impact of ESG on IG and HY bonds, our finding that crisis period outperformance is limited to the top HY Governance portfolio supports our assertion that agency problems (and more broadly credit risks) are crucial channels through which ESG influences bond returns.

6.4 Performance of alternative ESG indicators

Relative to the number of prominent ESG agencies and the importance of ESG ratings to investors and scholars, research into the quality of ESG scores and the transparency of those who provide them have been sparse and, according to some, inadequate (Windolph 2011, Hoepner et al. 2016, Chatterji et al. 2014, Drempetic et al., 2019). Such research is critical, because as Schäfer et al. (2004) and Khan, Serafeim, Yoon (2015) explain in detail, there are many issues inherent in providers' attempts to capture a topic as broad and as complex as sustainability within single ratings. La Torre et al. (2020) combine various ESG ratings with several other ESG indicators, including the opinions of financial analysts and other pieces of qualitative data, to construct a more holistic measure of a company's ESG profile.[120] While they catalogue the imperfections of existing ESG scoring systems, they note that scores, such as those from Thomson Reuters/Refinitiv, remain among the best (and often only) metrics for empirical researchers who seek to analyse the ESG-CFP relationship. Unfortunately, as Berg, Koelbel and Rigobon (2020) note, variability in ratings across providers means that the results of an empirical study are likely to be highly dependent on the researcher's choice of ESG score. The authors present a rigorous and comprehensive examination of the relationship between six of the largest ratings providers (including Refinitiv/Thomson Reuters), and find that the correlations between ratings are on average 0.54 and range from 0.38 to 0.71. They suggest that these variations not only pose challenges for empirical research, but that they also dilute any potential demand-based effects of ESG ratings on asset prices. Furthermore, the authors attribute ratings divergence to what, and how, providers measure ESG activities, which makes it particularly challenging for researchers to reconcile differences in ESG scores.

In an attempt to address some of these issues, we examine alternative ESG indicators. Specifically, we analyse the performance of portfolios formed on the year-over-year percentage change in ESG ratings (ESG momentum), a news-based ESG controversies score and lastly ESG ratings from a different provider (the Bloomberg ESG score). Although the former two measures do not address the

[119] Thomson Reuters claims to measure management's ability to integrate ESG policies, but we suggest that this effectively measures stakeholder engagement as well given that it is a critical component of successful ESG integration (Waddock and Graves, 1997, Gerard, 2018).

[120] The authors define ratings and scores separately. We do not carry over this distinction and use the two terms interchangeably.

issue of ratings divergence directly, they reflect information that researchers show is absent (or at least obscured) from the level of traditional ESG scores (Kim and Statman, 2012, Gregory and Whittaker, 2013). As a result, they provide us with a broader perspective on the drivers of the returns to ESG-based investment strategies.

6.4.1 Bloomberg ESG score

To combat the issues that Berg et al. (2020) raise it is critical that researchers re-test their results using different ESG ratings, especially in fixed income where the results of existing studies have, in many cases, not been replicated. Ideally, we would replicate our results with multiple ratings from agencies that employ varying methodologies, but given our limited access to data and our need for ratings with expansive coverage, we settle for Bloomberg's ESG scores. These ratings are popular among practitioners and encompass thousands of companies and 120 environmental, social and governance indicators. Additionally, unlike Thomson Reuters/Refinitiv, Bloomberg includes information from direct company contacts and sustainability assessments from third parties.[121] However, although even recent empirical studies rely on Bloomberg ESG scores as indicators of ESG performance (Ielasi et al., 2020), these scores are measures of ESG disclosure, not performance. Despite the distinctions between the two (Gold and Heikkurinen, 2017 and Drempetic et al., 2019), transparency remains a crucial component of sustainability (Mena and Palazzo, 2012) and there is some evidence of a positive relationship between disclosure and performance (Qian and Schaltegger, 2017). While tests with additional ESG scores would greatly enhance our analysis, we therefore assert that Bloomberg ESG scores are sufficient for our purposes.

Table 12 presents the cross-sectional characteristics of quintile portfolios formed using Bloomberg ESG scores and alpha statistics from regressions of Models 1 and 2 on ESG level, ESG momentum and size-, rating- and maturity-controlled portfolios. Our results do not change dramatically when we form portfolios on Bloomberg ESG scores, which in part reflects the high correlation we find between Bloomberg and TR scores (0.73).[122] In IG, the bottom portfolio remains the top performer based on the level of ESG scores, but is usurped by the top portfolio when we consider the score's year-over-year percentage change. When formed on the level and change in ESG ratings, the middle portfolios in HY display the best performances. Unfortunately, the economic and statistical significance of the Bloomberg portfolios' returns are much lower compared to those of our TR portfolios. Even when we construct ESG momentum portfolios with a 10% cut-off, we fail to register a result that is statistically significant at the 10% level. Thus, while on the one hand our results are somewhat robust, or at least not undermined, by the replacement of our primary ESG rating, they also highlight the need for caution in generalizing results based on ratings from a single provider.

6.4.2 ESG Momentum

Selecting securities based on how their associated ESG ratings change over time, or "ESG momentum", has become a popular investment strategy in the equities space (Dorfleitner et al., 2015, Nagy, 2016). We are unaware, however, of an academic study that applies such strategies to corporate bond portfolios. More than just a potentially profitable indicator, ESG momentum may also provide investors with a more effective method of assessing a firm's ESG performance. Gregory and Whittaker (2013) criticize the tendency of researchers to focus on the level, rather than the change, in ESG

[121] https://frameworkesg.com/bloomberg-esg-disclosure-scores-behind-the-terminal/
[122] ESG disclosure is, after all, a component of TR scores as well.

ratings. The authors show that firms adjust ESG policies infrequently, and that the largest valuation effects follow these policy shifts (Kim and Statman, 2012). They find a very high correlation between current and lagged KLD ESG scores (in excess of 0.9), and argue that this makes it difficult for researchers who use the level of ESG ratings to identify the valuation effects of ESG.[123] The correlation between an issuer's current and lagged TR score in our sample is 0.82, which, while not quite as high as those in Gregory and Whittaker (2013), raises the same questions regarding the accuracy with which base ESG scores capture changes in a firm's ESG performance. These issues, together with evidence in the existing literature that portfolios of stocks from companies with improving ESG ratings outperform both the market and other ESG equity portfolios (Agarwal and Ouaknin, 2019), lead us to examine ESG momentum in corporate bonds.

Table 13 shows the characteristics of quintile portfolios formed on the year-over-year percentage change in an issuer's ESG rating. In contrast to the trends we witness in Table 6 regarding the cross-sectional characteristics of portfolios sorted on the level of ESG ratings, we see very little variation in terms of average bond size, credit rating or maturity across ESG momentum quintile portfolios. It seems that ESG momentum may therefore provide a more neutral indicator of ESG performance, which does not suffer from some of the biases embedded in the level of ESG scores (e.g. the tendency for larger companies to receive higher scores regardless of sustainability performance (Drempetic et al. (2019)).[124] Another striking observation concerns the level of average ratings of ESG momentum portfolios. Top and bottom quintile portfolios in both IG and HY display average ESG scores that are markedly lower than those of the middle portfolios (between 0.5 and 1 standard deviation lower). This result is fairly intuitive given that scores are bounded (between 0 and 100), restricting the capacity for very high (low) scoring firms to continuously raise (lower) their ESG ratings.[125] In addition to controls on bond size, credit rating and maturity, we therefore implement our double sorting procedure with respect to the level of ESG ratings as well.

In IG, the performances of our ESG momentum portfolios diverge from those of our ESG level portfolios. Notably, both top and bottom momentum portfolios post distinctly higher average excess returns and Sharpe ratios compared to the middle portfolios (average Sharpe ratio of the top and bottom portfolios is 0.91 compared to an average of 0.72 for the middle portfolios). These results remain largely unchanged even after we control for differences in the level of ESG ratings (0.89 and 0.74).[126] Table 14, which displays the alpha statistics of ESG momentum quintile portfolios, shows that the top and bottom portfolios in IG produce statistically significant alphas of 0.66% and 0.60% based on Model 1. These alphas decline, however, using Model 2 (0.52% and 0.21% respectively), according to which neither intercept is significant at the 10% level. Even after controlling for bond size, credit rating and maturity the top ESG momentum portfolio maintains its performance advantage. At a 10% cut-off, the alpha of the top IG momentum portfolio increases further while that of the bottom portfolio decreases, resulting in a long-short portfolio that produces a positive and statistically significant alpha of 0.33%.

[123] They find that correlations are even higher for larger companies (0.94 and 0.95 respectively).

[124] This is of course assuming that bond size proxies for company size, which is certainly not always the case, especially in capital-light industries.

[125] It is also consistent with Gregory and Whittaker (2013) and Nagy et al. (2016) who find that firms that experience the largest changes in ESG scores are typically those whose scores are lower than average.

[126] The results do look different at a 10% cut-off, however, with the top (bottom) portfolio posting higher (lower) Sharpe ratios. To remain consistent with our earlier analyses we decide to carry on with the 20% cut-off rate. While it does not materially alter our conclusions, we reference our results with the 10% cut-off rate when appropriate.

The extant literature offers two interpretations of our results in IG. The first is along the lines of Gregory and Whittaker (2013) who argue that ESG momentum provides a clearer picture of the true relationship between ESG performance and financial returns. The authors show that stocks of companies with large ESG score upgrades or downgrades outperform stocks from companies whose scores remain relatively unchanged.[127] They suggest that this observation is consistent with theories of "strategic CSR" (a term coined by Baron, 2001), which assert that managers make decisions regarding ESG policy within the context of profit-maximization (Albuquerque et al., 2019). Since changes in a firm's ESG scores reflect strategic decisions to better align the interests of management with those of shareholders, investors reward both score upgrades and downgrades. More traditional explanations regarding the benefits to strong ESG practices may also apply, especially when we consider that the alpha of the bottom portfolio nearly vanishes with Model 2. Some of these benefits include a reduction in the cost of capital (Oikonomou et al., 2014), better access to financing (Amiraslani et al., 2019), higher asset value (largely through intangible asset creation, e.g. Edmans, 2011), improved resource efficiency (Hart, 1995) and a series of long-term, competitive advantages that accrue to firms with more supportive stakeholders (Porter and Kramer, 2007).

Our second interpretation concerns bondholders' perceptions of changes to a firm's ESG score. Derwall et al. (2011) suggest that markets underreact to positive ESG-related news, and Cui and Doherty (2020) show that, while ESG scores and stock prices exhibit little to no relationship in general, markets overreact to ESG-related controversies. In either case, that investors appear to misprice ESG momentum in the short term does not necessarily invalidate our earlier assertion that they properly value the effects of ESG on IG bonds in the long run. Moreover, investor preferences for securities from high ESG firms may differ from those for securities from firms that have improved their ESG scores, but which nonetheless remain at a low level. This could also explain why the existence of a shunned-stock effect is not as apparent in portfolios formed on ESG momentum as it is in standard ESG portfolios. In contrast to the trends we see in Tables 6 and 7, the performance of bottom quintile momentum portfolios in IG appears to have moderated over time.[128]

Unlike their IG counterparts, middle HY momentum portfolios display higher risk-adjusted compared to top and bottom quintiles (average Sharpe ratios of 0.82 compared to 0.62). Although not one of them generates a statistically significant intercept, there is a notable gap between the alphas of the bottom quintile (-1.45%) and the second quintile (1.32%). Once again, we propose that the greater relevance of agency frictions between shareholders and bondholders in HY may help to explain these divergent results. If ESG reflects management's efforts to align its interests with those of its shareholders, changes in ESG policy at a high-risk firm could prove detrimental to the firm's creditors (Cremars, 2007).[129] This is also consistent with SBRT, according to which the risks of asset substitution rise as firms spend more on ESG (Hoepner and Nilsson, 2020). Since HY firms are more sensitive to such risks, bondholders may be more likely to penalize them for dramatic changes in their ESG scores.

Also in contrast to IG, ESG level and ESG momentum portfolios behave similarly in HY. One notable deviation, however, concerns the underperformance of the bottom momentum quintile relative to the top. As a result, HY long-short momentum portfolios produce alphas between 0.76% and 2.44% (including double sorted portfolios, Table 13). This result is difficult to reconcile with the existing literature, which better explains the underperformance of top quintile portfolios. We suggest, that

[127] Kim and Statman (2012) arrive at a similar results with a different set of ESG scores.

[128] Ultimately, however, patterns in the time-series of ESG momentum returns are less clearly defined than they are in those of ESG level returns.

[129] As we describe in our literature review (Section 4), Cremars (2007) find a positive relationship between measures of shareholder control and bond yields.

with the exception of those related to asset substitution, agency risks can still account for the bottom portfolio's performance. As Amiraslani et al. (2019) note, bonds of firms that are closer to the default boundary are more sensitive to changes in CSR, and while strong CSR policies can benefit such firms, the costs of poor CSR are even more pronounced. The next section, which explores the influence of ESG controversies on bond returns, sheds more light on this phenomenon.

Table 12: Characteristics and performance of Bloomberg ESG portfolios

This table shows bond characteristics and performance statistics of our Bloomberg ESG-sorted portfolios for U.S. Investment Grade (IG) and U.S. High Yield (HY) corporate bonds over the period January 2015 to March 2021. The left hand side of the table shows long-only portfolios that go long per quintile of issuers' Bloomberg ESG scores, as well as long-short portfolios that go long the top 20% and short the bottom 20% of bonds based on issuers' Bloomberg ESG scores. We measure Credit Rating as the average across S&P, Moody's and Fitch ratings, Amt Out as the amount of debt outstanding, and corporate bond returns as annualized excess returns versus maturity-matched (Panel A). Panel B reports the annualized alpha of each portfolio with respect to two multi-factor models, Models 1 and 2, described in Section 6.2, as well as the models' adjusted R^2. *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence levels, respectively, of two-sided tests on whether the alphas are different from 0. *t*-statistics with Newey-West standard errors.

Panel A: Long-only portfolio bond characteristics

	ESG	Credit Rating	Amt Out	Yrs to Maturity	Duration	Excess Returns	Volatility	Sharpe ratio
				Investment Grade				
top 20%	56.79	7.28	762.2	8.06	6.05	2.96%	4.40%	0.672
20-40%	43.31	7.46	602.8	8.06	6.08	3.40%	4.91%	0.692
40-60%	32.75	7.70	619.6	8.57	6.29	3.73%	4.95%	0.753
60-80%	22.36	7.79	501.4	8.87	6.13	3.45%	4.33%	0.796
bot 80%	14.11	7.65	409.3	8.21	6.01	3.51%	4.75%	0.740
				High Yield				
top 20%	48.06	11.37	585.0	8.87	6.14	4.42%	8.15%	0.543
20-40%	28.24	11.82	453.1	8.87	5.99	5.96%	7.22%	0.825
40-60%	20.39	11.78	395.8	7.87	5.96	5.45%	7.20%	0.758
60-80%	16.02	12.15	456.9	7.99	6.18	5.65%	7.01%	0.807
bot 80%	12.43	11.80	485.7	7.86	5.81	5.40%	10.79%	0.500

Panel B: Performance statistics of long-only and long-short portfolios

	Investment Grade				*High Yield*			
	Model 1		Model 2		Model 1		Model 2	
	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2
top 20%	-0.27%	0.969	-0.11%	0.972	-1.71%	0.917	-1.94%**	0.917
	(-0.84)		(-0.39)		(-1.62)		(-2.03)	
20-40%	-0.42%	0.937	-0.36%	0.937	0.90%	0.886	1.33%**	0.905
	(-0.87)		(-0.67)		(1.41)		(2.09)	
40-60%	-0.26%	0.908	0.16%	0.909	0.73%	0.916	0.25%	0.924
	(-0.64)		(0.30)		(0.98)		(0.39)	
60-80%	0.40%	0.914	0.83%**	0.950	0.87%	0.846	0.90%	0.847
	(1.21)		(2.22)		(1.18)		(1.06)	
bot 80%	0.18%	0.844	0.99%*	0.903	-1.02%	0.915	-0.66%	0.919
	(0.49)		(1.77)		(-0.76)		(-0.47)	
Long/ Short	-0.46%	0.077	-1.10%	0.338	-0.69%**	0.041	1.27%***	0.129
	(-0.76)		(-1.54)		(-2.07)		(-2.98)	

Table 13: ESG Momentum Performance Statistics

This table shows the performance statistics of ESG Momentum-sorted portfolios for U.S. Investment Grade and U.S. High Yield corporate bonds over the period January 2015 to March 2021. We measure ESG Momentum as the year-over-year percentage change in each issuer's Thomson Reuters (TR) ESG score. Panel A shows the performance of long-only portfolios that select long positions based on quintiles of ESG Momentum each month. Panel B shows the performance of long-short portfolios that go long the top 20% portfolio and short the bottom 20%. We construct size-, credit- and duration-neutral portfolios by first creating long-short portfolios per quintile of debt outstanding, rating group (AAA/AA+/AA/AA-/A+, A/A-, BBB+, BBB, BBB-, BB+/BB, BB-/B+/B), or quintile of [years to maturity / duration] and then combining all groups to form the final portfolio. The table reports the annualized alpha of each portfolio with respect to two multi-factor models, Models 1 and 2, described in Section X, as well as the models' adjusted R^2. *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence levels, respectively, of two-sided tests on whether the alphas are different from 0. *t*-statistics with Newey-West standard errors. We measure corporate bond returns as excess returns versus maturity-matched Treasuries.

	Investment Grade				*High Yield*			
	Model 1		**Model 2**		**Model 1**		**Model 2**	
	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2
Panel A: ESG Momentum Quintiles								
top 20%	0.66%*	0.923	0.52%	0.967	-0.34%	0.914	-0.10%	0.937
	(1.98)		(1.52)		(-0.38)		(-0.11)	
20-40%	0.07%	0.875	0.53%	0.893	1.08%	0.928	1.32%	0.939
	(0.09)		(0.60)		(1.02)		(1.38)	
40-60%	-0.03%	0.926	-0.39%	0.951	0.72%	0.824	0.76%	0.856
	(-0.08)		(-1.06)		(0.70)		(0.80)	
60-80%	0.01%	0.933	0.05%	0.957	0.79%	0.901	0.13%	0.923
	(0.02)		(0.17)		(1.06)		(0.23)	
bot 80%	0.60%**	0.944	0.21%	0.971	-0.96%	0.920	-1.45%	0.945
	(2.30)		(1.50)		(-0.71)		(-1.33)	
Panel B: ESG Momentum Long Short								
base	-0.06%	0.089	0.31%	0.159	0.56%	0.003	1.32%	0.396
	(-0.21)		(-0.16)		(0.33)		(0.91)	
size neutral	-0.15%*	0.412	0.38%*	0.536	1.87%	0.307	2.44%	0.404
	(-1.75)		(1.90)		(1.08)		(1.24)	
credit neutral	-0.23%**	0.403	-0.29%	0.515	1.26%***	0.356	0.76%***	0.383
	(-2.03)		(-1.52)		(3.43)		(2.43)	
duration neutral	0.54%*	-0.036	0.23%*	0.167	1.03%	0.027	2.27%*	0.311
	(1.88)		(1.74)		(0.75)		(1.85)	

6.4.3 ESG Controversies

Another advantage of the TR score is the newly implemented "Controversies" score. Thomson Reuters assesses companies on 23 topics concerning ESG-related controversies (the list of which is available in the appendix of Thomson Reuters, 2019), and provides a score that reflects the intensity of negative ESG events reported in major media publications. The higher the Controversies score, the larger the downgrade in a company's ESG rating. To construct a separate Controversies score, we measure the magnitude of this downward adjustment in percentage terms (which we will refer to as the CP-score).[130] Since most firms do not experience ESG controversies on a regular basis, we cannot compare the cross-sectional characteristics of quintile portfolios sorted on CP-scores. Instead, we compare the returns of a portfolio comprised of bonds without CP-scores (which we call the "CP-free" portfolio) to portfolios of bonds with CP-scores. Regarding the former, we create three portfolios based on different CP-score thresholds (0%, 5% and 10%), such that our CP-0 portfolio contains all bonds with CP-scores, and our CP-10 portfolio only contains bonds from firms whose controversies result in at least a 10% score reduction. We also form long-short portfolios, which fund a position in the CP-free portfolio with a position in one of the three CP portfolios. To our knowledge, we are the second study to construct portfolios using TR's Controversies score (the first being Dorfleitner et al., 2020), and the first to do so with corporate bonds.

In conducting our standard robustness checks described in Section 6.3c, we find that, for the first time in our analysis, splitting our sample along one of our characteristic measures (in this case amount of notional outstanding) drastically changes our results (summarized in Table 14 below). This is not so in IG, however, where the CP-free portfolio generates a small but statistically significant alpha of 0.33% and the three CP portfolios produce insignificant alphas. Splitting our sample according to bond size does not alter our rather mild results. In HY, the CP-free portfolio produces an economically and statistically insignificant alpha. CP0, on the other hand, generates an alpha of -0.98% (-1.03), which increases to -1.85% (-1.76) as we increase the score threshold to 5% (CP5). Although these results are robust to controls on credit quality and maturity, they appear to be highly sensitive to changes in bond size. To demonstrate this we divide our sample in two based on bond size and create two new sets of CP portfolios. As we see in Table 14, the alphas on all three of our CP portfolios comprised of the smallest 50% of bonds fall sharply to statistically significant values of -4.13%, -5.99% and -5.43%. In other words, investors who had funded positions in the CP-free portfolio with short positions in any one of the CP portfolios would have reaped sizeable profits over our sample period. This is not the case among the largest 50% of bonds, however, as the performances of the CP portfolios are offset by the comparable alphas of the CP-free portfolios. Moreover, while among the smallest bonds there appears to be a negative relationship between the intensity of ESG-related controversies and financial performance, in the large bond sample the alphas (though not statistically significant) rise from CP0 (0.97% (0.92)) to CP10 (1.63% (1.01)).

These results further support our hypothesis concerning the role of credit quality (or, more specifically, of being classified as HY) in mediating the influence of ESG on bond returns. That we see negative alphas in HY is consistent with the results of Hoepner and Nilsson (2020), as well as with our initial results, which show that across our sample period, portfolios of bonds with low ESG scores underperform portfolios with middling scores. If it is through credit risk that ESG influences bond returns (see Halling et al. (2020)) then bonds that are closer to default should be more sensitive to both positive and negative changes in ESG. As Bauer and Hann (2014) note, ESG-related controversies can be

[130] Or the percentage difference between the combined ESG score and the combined ESG score with the controversies overlay.

large enough that firms consider filing for bankruptcy in order to avoid them (or else consider other tactics that disproportionately harm to bondholders). Additionally, as we describe earlier, trading frictions may cause investors in HY to discount ESG strengths and controversies less efficiently than do investors in IG (Jostova, 2013). That smaller bonds are less liquid and less efficiently priced (Sarig and Warga, 1989), could therefore help to explain the disparity between the results in portfolios of small HY bonds compared to those of large HY and IG bonds. Alternatively, our results may reflect the presence of mechanisms that are similar to those Dorfleitner et al. (2020) describe concerning firm size and ESG. The authors show that markets react more sharply to ESG-related controversies in smaller companies with fewer resources. Due to a higher rate of return and, presumably, fewer slack resources, it appears that small bonds in HY are the most sensitive to ESG controversies.

Table 14: ESG Controversy Performance Statistics

This table shows the performance statistics of ESG Controversy-sorted portfolios for U.S. Investment Grade and U.S. High Yield corporate bonds over the period January 2015 to March 2021. Our ESG Controversy score represents the percentage difference between each issuer's Thomson Reuters (TR) ESG Combined score and its TR ESG Combined score with controversies overlay. We organize the table in three primary columns to compare our results when we exclude bonds based on amount of debt outstanding (Size): size-neutral includes all bonds in our sample; Small Bonds includes the 50% smallest bonds; Large Bonds includes only the 50% largest bonds. On the left hand side of the table, Panel A shows the performance of long-only portfolios of bonds from issuers without ESG Controversies. Panels B, C, and D show the performance of long-only portfolios of bonds from issuers with an ESG Controversy score above 0%, 5%, and 10%, respectively, as well as long-short portfolios that go long the controversies-free portfolio and short the portfolios of bonds from issuers with ESG Controversies above 0%, 5%, and 10%, respectively. The table reports the annualized alpha of each portfolio with respect to two multi-factor models (Models 1 and 2) described in Section X, as well as the models' adjusted R^2. *, **, and *** indicate statistical significance at the 90%, 95%, and 99% confidence levels, respectively, of two-sided tests on whether the alphas are different from 0 (*t*-tests with Newey-West standard errors). We measure corporate bond returns as excess returns versus maturity-matched Treasuries.

	Size Neutral				**Small Bonds**				**Large Bonds**			
	Model 1		Model 2		Model 1		Model 2		Model 1		Model 2	
	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2	Alpha (*t*-val)	adj R^2
					Investment Grade							
Panel A: Long below 0%												
Long	0.19%**	0.997	0.33%***	0.997	0.36%	0.963	0.17%	0.987	0.06%	0.960	0.55%**	0.983
	(2.04)		(3.41)		(1.23)		(0.91)		(0.17)		(2.16)	
Panel B: Short above 0%												
Short	0.14%	0.979	-0.09%	0.981	0.40%	0.944	-0.08%	0.962	-0.24%	0.931	-0.29%	0.964
	(0.51)		(-0.30)		(0.88)		(-0.20)		(-0.48)		(-0.76)	
Long/Short	0.05%	-0.010	0.42%	0.097	-0.04%	0.081	0.26%	0.194	0.30%	0.250	0.85%*	0.351
	(0.14)		(1.16)		(-0.09)		(0.58)		(0.61)		(1.78)	
Panel C: Short above 5%												
Short	0.29%	0.972	0.08%	0.973	0.53%	0.933	0.08%	0.955	0.01%	0.929	-0.03%	0.955
	(0.87)		(0.24)		(1.08)		(0.17)		(0.03)		(-0.08)	
Long/Short	-0.09%	0.060	0.25%	0.100	-0.17%	0.030	0.10%	0.149	0.05%	0.332	0.58%	0.376
	(-0.24)		(0.64)		(-0.36)		(0.21)		(0.09)		(1.08)	
Panel D: Short above 10%												
Short	0.46%	0.966	0.27%	0.967	0.74%	0.925	0.28%	0.948	0.11%	0.927	0.14%	0.953
	(1.22)		(0.72)		(1.45)		(0.57)		(0.21)		(0.36)	
Long/Short	-0.26%	0.105	0.06%	0.110	-0.38%	0.006	-0.10%	0.064	-0.05%	0.353	0.42%	0.394
	(-0.62)		(0.15)		(-0.78)		(-0.21)		(-0.08)		(0.75)	
					High Yield							
Long below 0%												
Long	0.39%	0.985	0.08%	0.986	-0.03%	0.962	-0.86%	0.968	1.02%	0.943	1.29%*	0.973
	(0.87)		(0.16)		(-0.05)		(-1.23)		(1.34)		(1.79)	
Short above 0%												
Short	-0.86%	0.918	-0.98%	0.927	-3.80%	0.749	-4.13%**	0.781	0.85%	0.903	0.97%	0.911
	(-0.83)		(-1.03)		(-1.64)		(-2.03)		(0.87)		(0.92)	
Long/Short	1.25%	0.221	1.05%	0.341	3.76%	0.069	3.27%	0.152	0.17%	0.257	0.32%	0.366
	-0.9500		(0.91)		(1.53)		(1.58)		(0.14)		(0.27)	
Short above 5%												
Short	-1.58%	0.908	-1.85%*	0.924	-5.50%**	0.690	-5.99%**	0.725	1.17%	0.881	1.15%	0.893
	(-1.37)		(-1.76)		(-2.25)		(-2.61)		(0.96)		(0.88)	
Long/Short	1.97%	0.210	1.93%	0.385	5.47%**	-0.001	5.13%**	0.069	-0.15%	0.290	0.14%	0.427
	(1.37)		(1.54)		(2.13)		(2.21)		(-0.11)		(0.10)	
Short above 10%												
Short	-1.28%	0.901	-1.42%	0.919	-5.01%*	0.677	-5.43%**	0.708	1.51%	0.852	1.63%	0.869
	(-1.02)		(-1.30)		(-1.97)		(-2.30)		(1.01)		(1.01)	
Long/Short	1.67%	0.219	1.50%	0.400	4.97%*	0.005	4.57%*	0.060	-0.49%	0.268	-0.34%	0.404
	(1.10)		(1.18)		(1.87)		(1.93)		(-0.29)		(-0.20)	

6.5 Final results

To conclude our empirical analysis, we examine the "investable" returns to popular ESG investment strategies, as well as those of our most profitable ESG portfolios from previous sections. We define investable returns as those to a long-only portfolio net of transaction costs (which we assign to each bond based on estimates of bid-ask spreads and trading costs in Chen et al. (2007)).[131] We incorporate ethical screening and best-in-class strategies, which according to USSIF (2020), account for the highest percentage of sustainable assets under management.[132] To implement the latter we form portfolios that exclude all bonds from controversial businesses (the bond equivalent of so-called "sin" stocks, Hörter (2017)). We base our exclusion criteria on the UNPRI's (2021) list of the most popularly excluded industries, which include tobacco, firearms, alcohol, and gambling. The best-in-class strategy (sometimes to referred to as positive screening) involves the identification of the top companies within each industry in regards to pre-defined ESG characteristics.[133] Since TR scores are constructed to be industry neutral, our top quintile portfolios essentially represent a form of the best-in-class strategy.[134] We therefore take a slightly unconventional approach to build our best-in-class portfolios, such that we retain all bonds in our sample and adopt a rank-weighting scheme similar to that of Dorfleitner et al. (2020). This means that, within each industry subgroup, we weight each bond according to its TR score (i.e. a bond with a score of 80 receives a weight that is twice that of a bond with a score of 40). Finally, we follow Hoepner and Nilsson (2020) and append industry dummies to our most successful multifactor model (Model 2), which creates Model 4 (see Section 6.2). Our dummies represent the orthogonalized returns to portfolios of bonds grouped according to their ICB classifications. Table 15 displays the return statistics and average ESG scores of our ESG strategy portfolios and our top performing ESG quintile portfolios.

Consistent with the literature concerning the application of prominent ESG strategies to equity portfolios (Breedt et al., 2018), we find that such strategies would have enabled investors to improve a bond portfolio's sustainability profile without sacrificing its financial performance. Our IG and HY best-in-class portfolios perform on par with their respective market portfolios but boast average ESG scores that are approximately 20% larger.[135] Although exclusionary strategies do not engender a similar bump in average ESG scores, we argue that they reflect alternative sustainability goals that may not be well reflected in ESG ratings (especially those that are normalized at an industry level). Our exclusionary portfolio in HY posts a respectable, and statistically significant, alpha of 0.92% according to Model 1, but this alpha disappears entirely when we assess the portfolio against Model 4. Our results are similar to those of Li and Zhang (2016), who apply sin screens to portfolios of U.S. corporate bonds from 2004 to 2015 and observe no significant impact on bond returns.

Our best quintile portfolios continue to perform admirably after we account for transaction costs and industry effects. With the exception of ESG momentum, all of our bottom quintile portfolios

[131] The use of transaction-based prices would allow us to calculate transaction costs more accurately. Our current approach yields estimates that are only slightly better than proportional transaction costs.
[132] While common place in the equities space, these strategies have received comparatively little attention in the corporate bond market.
[133] For one definition of the best-in-class strategy see https://www.robeco.com/ch/en/key-strengths/sustainable-investing/glossary/best-in-class.html
[134] Since we do not have data on all the companies that the TR score covers, our quintile portfolios are not completely industry neutral.
[135] According to Model 4, the HY best-in-class portfolio posts a small negative alpha, but the intercept is not statistically significant.

in IG produce statistically significant alphas of ~0.6% with Model 4. This strong financial performance, however, comes at the cost of a portfolio's average ESG score, which declines 20% to 35%. The alphas of our HY portfolios are more model dependent. Based on Model 1, our second quintile Combined, Environmental, and Governance portfolios display relatively strong alphas of 1.74% (2.21), 1.68% (2.30), and 2.15% (2.74) respectively. These fall sharply with Model 4, however, and only our Combined and Governance portfolios retain statistically significant alphas (0.52% (1.73) and 1.68% (1.77)).[136] Unlike our IG portfolios, our HY portfolios exhibit average ESG scores that are between 10% and 20% above the mean. It is also notable that, although our quintile portfolios represent only a fraction of the market portfolio, each second quintile in HY exhibits less negative skewness and a lower kurtosis. Jagannathan et al. (2017) find a similar trend with respect to higher moments of the return distributions of ESG-integrated equity portfolios, which they attribute to a reduction in various downside risks associated with poor ESG practices.

[136] Though not statistically significant at a 10% level, the alphas on our Environmental and Momentum portfolios are at least worth noting (both are about 1% with t-stats of 1.32 and 1.52).

Table 15: ESG strategies performance statistics

This table shows the performance statistics of various ESG-related strategies for U.S. Investment Grade (IG) and U.S. High Yield (HY) corporate bonds over the period January 2015 to March 2021. In both IG and HY, we evaluate a best-in-class strategy that weights bond returns by issuers' Thomson Reuters (TR) ESG Combined Scores and a strategy that excludes companies in the tobacco, alcohol, gambling and firearms industries. In IG, we assess strategies that go long the top 20% of bonds based on year-over-year percentage change in issuers' ESG Combined scores (ESG Momentum) and go long the bottom 20% of bonds based on ESG Combined scores, Environment scores, Social scores, and Governance scores, respectively. In HY, we assess strategies that go long bonds that fall within the 20%-40% percentile [second highest quintile] of ESG Combined Scores, Environment scores, Social scores, Governance scores, and ESG Momentum scores, respectively. We calculate returns as the average of the portfolios constructed from month $t-1$ to month t. We measure corporate bond returns as excess returns versus maturity-matched Treasuries. Panel A shows statistics on the returns series for each strategy, as well as the corporate bond market (DEF). Panel B shows alpha statistics with respect to a multi-factor model (RMRF, SMB, HML, MOM, TERM and DEF). Panel C shows alpha statistics with respect to a multi-factor model, Model 4, which includes returns of industry portfolios (orthogonalized (w.r.t to the market portfolio) (see description in Section 6). We annualize all alphas and report the models' adjusted R^2. *,**, and *** indicate statistical significance at the 90%, 95%, and 99% confidence levels, respectively, of two-sided tests on whether the alphas are different from 0. t-statistics with Newey-West standard errors.

	Investment Grade							
	Equal	Best In Class	Exclusionary	Top ESG Momentum	Bottom ESG	Bottom Env	Bottom Gov	Bottom Soc
Panel A: Returns Series statistics								
mean	3.63%	3.75%	3.69%	4.55%	3.80%	3.80%	3.69%	3.97%
volatility	4.60%	4.64%	4.47%	4.99%	4.48%	4.09%	4.16%	4.62%
Sharpe ratio	0.79	0.81	0.82	0.91	0.85	0.93	0.89	0.86
skewness	-1.00	-0.88	-0.78	-1.01	-1.01	-0.71	-0.94	-0.78
kurtosis	4.30	3.83	3.36	6.15	4.86	3.30	3.63	3.09
ESG mean	61.56	71.56	63.41	56.83	35.92	39.29	49.83	39.11
Panel B: Model 1 Results								
alpha		0.10%	0.16%	0.54%	0.46%*	0.64%***	0.77%	0.51%**
t-value		1.080	0.940	1.578	1.824	2.798	1.631	2.171
adj R^2		0.997	0.994	0.963	0.959	0.960	0.949	0.967
Panel C: Model 4 Results								
alpha		0.13%	0.06%*	0.32%	0.65%**	0.56%**	0.62%**	0.67%**
t-value		1.540	1.909	0.977	2.246	2.565	2.264	2.045
adj R^2		0.998	0.999	0.977	0.983	0.985	0.980	0.974

	High Yield							
	Equal	Best In Class	Exclusionary	2nd Best ESG	2nd Best Env	2nd Best Gov	2nd Best Soc	2nd ESG Momentum
Panel A: Returns Series statistics								
mean	5.73%	5.76%	5.70%	6.99%	6.94%	7.38%	6.32%	7.94%
volatility	8.12%	7.92%	6.59%	8.48%	7.71%	7.63%	9.19%	9.23%
Sharpe ratio	0.70	0.73	0.87	0.82	0.90	0.97	0.69	0.86
skewness	-3.82	-3.42	-3.46	-2.95	-3.46	-2.94	-3.02	-3.07
kurtosis	26.15	23.48	22.23	22.29	23.62	19.25	22.66	22.28
ESG mean	52.28	60.28	52.97	62.88	61.18	59.56	60.84	56.12
Panel B: Model 1 Results								
alpha		-0.02%	0.92%**	1.74%**	1.68%**	2.15%***	0.29%	1.08%
t-value		-0.048	2.357	2.213	2.307	2.737	0.265	1.018
Panel C: Model 4 Results								
alpha		-0.63%	-0.09%	0.52%*	1.00%	1.68%*	-0.22%	1.07%
t-value		-1.659	-0.709	1.727	1.319	1.767	-0.207	1.522
adj R^2		0.988	0.998	0.938	0.953	0.933	0.929	0.947

7. Conclusion

In contrast to the extensive literature on the time-series and cross-sectional determinants of stock returns, as well as the burgeoning research on the effects of ESG, there is surprisingly little research in either area with respect to bond returns. To address these gaps, this paper analyzes the performance of bond-specific factor mimicking portfolios and ESG-sorted corporate bond portfolios. The portfolios we construct to capture well-established risk factors generate abnormal returns, and also help us to measure the performance of ESG-sorted portfolios. We show that certain ESG portfolios generate returns that our multifactor models fail to explain, and that these returns are robust to both controls for prominent bond characteristics, and to the incorporation of alternative ESG indicators and transaction costs. Lastly, we demonstrate that popular ESG investment strategies can be implemented in bond portfolios without sacrificing financial performance. To conduct our empirical analysis, we collect data from Bloomberg and Thomson Reuters Eikon on corporate bonds from 1950 unique, U.S. issuers over the period January, 2015 to March, 2021. We use filters from Haesen et al. (2013) and Israel et al. (2018) to select a representative bond from each issuer each month, which allows us to construct a cross-sectionally comparable panel of liquid corporate bonds. In line with evidence of market segmentation along the BBB-/BB+ divide (Ambastha et al., 2010, Chen et al., 2014) and the tendency of investors, regulators, and index providers to treat IG and HY bonds as separate asset classes (Houweling and van Zundert, 2017), we split our sample between these two ratings categories.

In the first part of our analysis, we use the sorted portfolio approach of Fama and French (1992) to construct bond portfolios that proxy for well-known factors. We borrow factor definitions from recent empirical studies, including Jostova et al. (2013), Bai et al. (2018), and Houweling and van Zundert (2017), and find the performances of our long-short portfolios to be remarkably consistent with those of the latter. We show that allocations to Value, Size, and Momentum would have enabled investors to generate abnormal risk-adjusted returns during our sample period, but that these returns vary across IG and HY markets. Most notably, while we present evidence of a strong momentum effect in HY (Momentum portfolio produces an alpha of 8.82%), we show a comparable reversal effect in IG (alpha of -4.24%). In response to criticism regarding researchers' use of unrealistic zero-cost factor-mimicking portfolios (Blitz et al., 2018), we also test long-only variants that incorporate transaction cost estimates. While the short legs of our factors portfolios artificially inflate the returns achievable in practice, we find that investable versions of these portfolios still yield positive and statistically significant alphas. Lastly, we use factor spanning tests from Huberman and Kubel (1987) and the orthogonalization procedure of Elton et al. (1993) to combine our factor portfolios into alternative multifactor models with which to explain the performance of our ESG-sorted portfolios. Due to a lack of alternatives, researchers rely on a combination of equity factors and bond indices to evaluate the performance of bond funds and synthetic portfolios (Henke, 2016, Pereira, 2019, Hoepner and Nilsson, 2020). In line with recent studies that challenges the effectiveness of such models (Bai et al., 2018, Bali et al., 2019), we find that our alternative multifactor models outperform our conventional ones.

The second part of our empirical analysis focuses on the performance of ESG-sorted corporate bond portfolios. Like Hoepner and Nilsson (2020) and Dorfleitner et al. (2020), we explore the ways in which the patterns we observe in the returns of our portfolios align with theoretical frameworks concerning the relationship between ESG ratings and bond risks and returns. The results of our IG and HY portfolios support two separate sets of explanations regarding the role of ESG in bond returns. In IG, bottom quintile portfolios based on Environmental, Social, Governance, and Combined scores produce statistically significant alphas that range from 0.6% to 1%. With the exception of our second

Environmental quintile, which generates a robust alpha comparable to that of its respective bottom quintile (0.8%), no other quintile portfolio is associated with abnormal financial performance. The time-series variation we observe in the returns of these portfolios aligns with the predictions of both the shunned-stock and errors-in-expectations hypotheses. In support of the former, which states that investors neglect bonds from poorly rated issuers and therefore push up their expected returns, we find that the alphas of our bottom quintile portfolios generally increase over our sample period. Moreover, if the market previously rewarded bonds from top-rated issuers as some researchers suggest (Derwall and Koedjik, 2009, Polbennikov et al., 2016), our observation that not one of our top quintile portfolios produces a positive and statistically significant alpha suggests that investors have learned to properly discount the benefits of ESG management over time. Remarkably, when we form quintile portfolios based on ESG momentum, both the top and bottom portfolios produce positive alphas. This suggests that investors price changes in ESG scores, especially pronounced ones, differently than they do the level of ESG scores, and/or that ESG momentum captures different aspects of ESG performance and its impact on firm and bond fundamentals (Kim and Statman, 2012, Gregory and Whittaker, 2013).

In HY, we show that portfolios of both the best and worst rated bonds tend to underperform those comprised of bonds with middling ESG ratings. This is most apparent in the differences between the alphas of top and second quintile portfolios formed on Combined, Environmental and Social scores, which reach 3.7%, 5.4%, and 1.4% respectively. We argue that the greater relevance of credit and agency risks can help explain the disparities we observe in the performances of quintile portfolios in HY compared to those in IG. For example, risks of asset substitution, cash diversion, and leveraged acquisition are more likely to materialize in firms that are closer to the default boundary (Amiraslani et al., 2019). As fixed claimants do not benefit equally from investments into ESG compared to residual claimants (such as shareholders), risks of asset substitution rise alongside ESG spending (Hoepner and Nilsson, 2020). Additionally, higher required rates of returns and fewer slack resources make investments into ESG more costly for HY issuers in general. Just as Dorfleitner et al. (2020) find that the market penalizes the equities of small companies for high levels of ESG spending, so too do investors appear to discriminate between HY and IG bonds with respect to ESG. However, since increased credit and agency risks are also associated with poor ESG practices, the lowest rated HY portfolios fail to perform like their IG counterparts. We assert that we see the benefits of ESG, such as better access to financing (D. Cheng et al., 2014), fewer agency frictions (Ferrell et al., 2016), and decreased regulatory (Bauer and Hann, 2014), litigation (Hong and Liskovich, 2016), and reputational risks, manifest in the performances of our second quintile portfolios, which are less likely to face the drawbacks associated with excessive ESG spending.

We make three additional observations that support our conclusions, but which require future research beyond the scope of this paper. First, we find that while our top HY Governance quintile produces a negative alpha during the period from 2015 to 2020 (-1.94%), it posts a strong, positive alpha of 2.7% from 2019 to 2021 (and during the Covid-19 market crash). As a result, the top Governance quintile is the only top quintile portfolio in HY not to underperform across our full sample. This is consistent with Amiraslani et al. (2019), who show that strong CSP track records helped insulate the bonds of vulnerable issuers (those with fewer tangible assets and lower credit ratings) from the effects of the GFC.[137] Like the authors, we suggest that Governance proxies for trust between management and bondholders, and that the benefits to strong Governance are more likely to manifest during periods when overall trust is low, as well as in bonds from firms that are more exposed to such risks (i.e. HY issuers). Unfortunately, the window of market turmoil following the spread of Covid-19 is too

[137] On average, such issuers benefitted from lower bond yields and a reduced likelihood of default (Amiraslani et al., 2019).

narrow for us to make strong inferences. With a sample that spans multiple phases of elevated volatility, future researchers can more effectively investigate the role of Governance (and ESG more broadly) in the performance of bond portfolios during crisis periods. Our results also highlight the importance that researchers examine component ESG indicators separately. Environmental, Social, and Governance sub-scores, however, still encompass a broad range of sustainability activities, some of which may have potentially conflicting effects on bond returns.[138] It is therefore critical that, as many have with equities, future researchers explore more granular ESG metrics with respect to corporate bonds (Hoepner et al., 2016).

Based on evidence that investors react differently to negative ESG-related news than they do to the level of ESG ratings (De Franco, 2020, Dorfleitner et al., 2020), we use the TR Controversies score to compare the performance of portfolios of bonds from issuers without ESG-related controversies to those with controversies of varying intensities. Although our initial analysis reveals no noteworthy trends, this changes after we split our sample according to bond size. Portfolios of the smallest 50% of HY bonds from issuers with ESG controversies significantly underperform our benchmarks (statistically significant alphas between 4% - 6%). To the extent that, within credit rating categories, bond size is associated with firm size, this result underscores how an issuer's vulnerability mediates the impact of negative ESG news on bond returns. Future researchers who incorporate data on a firm's fundamentals, especially balance sheet indicators, will be able to explore broader definitions of vulnerability. Moreover, the inclusion of equity data will allow researchers to directly compare the impact of ESG on equities and bonds from the same issuer, an analysis that Bektic (2018) conducts in regards to equity market factor definitions, but which no study has done with ESG.[139]

Lastly, our analysis of the factor exposures that underlie the returns of our ESG quintile portfolios provides more insight into the market's potentially asymmetric treatment of IG and HY bonds. Most notably, we detect robust positive (negative) coefficients on Value for top (bottom) ESG quintiles in HY. Although most studies do not separate IG and HY bonds, this result is unexpected in the context of the existing literature, which suggests that, all else equal, high ESG scores are associated with lower bond yields (Oikonomou et al., 2014, Polbennikov et a., 2016, Salvi et al., 2019, Halling, 2020). We see clearer evidence of this in our IG quintiles, as top (bottom) portfolios load negatively (positively) on Value (especially those formed on the Governance sub-score). Regressions of ESG scores on bond yields would enable researchers to compare how the market values ESG in IG and HY more effectively. From the perspective of portfolio management, however, it is important to understand how ESG relates to established factors, especially as investors increasingly extend factor-based strategies from equities to bonds (Slimane et al., 2020).

The broad scope of this paper's investigation, which we argue is appropriate given the nascent state of the literature, leads to many potential directions for future research. One of the least exciting, but most important, is the refinement and replication of our quantitative analysis. Additional tests with high quality data are crucial to address data mining concerns, especially since studies tend to rely on the same data (MacKinlay, 1995, Bektic, 2018). Furthermore, although the sorted portfolio method we use reduces measurement error and enhances the validity of our tests, it often leads to an erroneous rejection of the null hypothesis that the asset pricing model is true (Lo and MacKinlay, 1990). It would also be beneficial to examine the results of our analyses with transaction-based prices, which

[138] For example, previous research demonstrates that Governance measures that strengthen shareholder influence are positively related to bond yields (Cremars, 2007), while those related to ESG-integration and management exhibit a negative relationship (Salvi et al., 2019).

[139] While previous studies have compared the performance of equities and bonds with respect to well-established factor definitions (Bektic, 2018), researchers have yet to do so with ESG.

may improve the explanatory power of test statistics designed to detect abnormal bond returns (Bessembinder et al., 2006, 2009). In addition, transaction data would enable researchers to explore a liquidity premium. Although previous studies demonstrate the capacity for liquidity factors to explain cross-sectional differences in credit spreads (Longstaff et al., 2005, Lin et al., 2011), not one study has investigated the relationship between these factors and ESG criteria (Bektic, 2018).[140] Lastly, an examination of ESG and bond returns in an international context would present researchers with more ammunition against data mining arguments, assist investors who manage increasingly global portfolios, and open opportunities to study the role of investor preferences, market structure, and the regulatory environment (Menz, 2010, Leite and Cortez, 2016, Bektic, 2018, Badia and Cortez, 2020).

This paper has only started to delve into the implications of ESG on the management of corporate bond portfolios. Compared to ESG bond funds, our synthetic portfolios allow us to estimate the effects of ESG integration more precisely. Bond funds, however, provide a more realistic set of returns, and as ESG ETFs and mutual funds continue to proliferate, researchers should revisit and extend the work of Henke (2016) and Leite and Cortez (2016).[141] Moreover, the best-in-class and exclusionary strategies we examine can be formulated in many different ways, and there are a range of theme-based strategies that employ varying conceptualizations of sustainability. Additionally, investigations of the growing universe of green bonds are not only important to allay investor concerns about the new market (ICMA, 2018), but also provide researchers novel opportunities to isolate the effects of the “green” label (Zerbib, 2017). Lastly, our paper addresses ESG’s role in the performance of bond portfolios from the perspective of active management. Passive management techniques remain largely unexplored with respect to ESG bond portfolios (Slimane et al., 2020), and in order to close the gap between the percentage of investors who consider the ESG profile of their bond portfolios and those who do so with their equity portfolios, future research will need to address both areas.

[140] Some researchers have instead found ways to minimize the influence of liquidity, such as Barth et al. (2020) and Halling (2021) who examine the effects of ESG on CDS spreads and primary issue spreads respectively.

[141] The authors analyze the performance of SRI bond funds in the U.S. and Europe respectively. As ESG funds become more distinct and transparent, future researchers will be able to avoid some of the challenges that Henke (2016) and Leite and Cortez (2016) encountered (i.e. Henke (2016) estimates that at least 1/3 of the SRI funds in his sample are in fact conventional funds in disguise).

Bibliography

1. Aboody, D. and B. Lev, “The Value Relevance of Intangibles: The Case of Software Capitalization,” *Journal of Accounting Research* 36, Studies on Enhancing the financial Reporting Model (1998): 161-191.
2. Adamska, A., T. J. Dabrowski, and A. Grygiel-Tomaszewska, “The Resource-Based View of Stakeholder Theory: Which Better Explains the Relationship between Corporate Social Responsibility and Financial Performance?” *Eurasian Journal of Business and Management* 4(2), (2016): 1-16.
3. Albuquerque, R. A., Y. J. Koskinen, S. Yang, and C. Zhang, "Love in the time of COVID-19: The resiliency of environmental and social stocks," *Working paper* (2020): https://ssrn.com/abstract=3583611.
4. Albuquerque, R. A., Y. Koskinen, and C. Zhang, "Corporate Social Responsibility and Firm Risk: Theory and Empirical Evidence," *European Corporate Governance Institute (ECGI) - Finance Working Paper No. 359/2013* (2018): https://ssrn.com/abstract=1961971. http://dx.doi.org/10.2139/ssrn.1961971.
5. Alessandrini, F.; Jondeau, E., “ESG investing: From sin stocks to smart Beta,” *Swiss Financ. Inst. Res. Pap*. (2019): 19–16. Available online: https://ssrn.com/abstract=3357395
6. Ambastha, M., A. Ben Dor, L. Dynkin, J. Hyman, and V. Konstantinovsky, "Empirical Duration of Corporate Bonds and Credit Market Segmentation," The Journal of Fixed Income, 20(1) (2010): p. 5 – 27.
7. Amel-Zadeh, A. and Serafeim, G., “Why and how investors use ESG information: evidence from a global survey”, *Financial Analysts Journal* 74(3) (2018):87-103.
8. Amiraslani, H., K. V. Lins, H. Servaes, A. Tamayo, "Trust, Social Capital, and the Bond Market Benefits of ESG Performance," *Finance Working Paper* 535/2017 (2021).
9. Anand, A., C. Jotikasthira, K. Venkataraman, “Mutual Fund Trading Style and Bond Market Fragility,” *The Review of financial Studies*, hhaa120, (2020).
10. Anderson, R.W., and S. Sundaresan, “Design and Valuation of Dept Contracts*,” Review of Financial Studies*, 9(1) (1996): 37-68.
11. Andrews, S. and T. Vunaki. “The Evolution of Responsible Investment: An Analysis of Advanced Signatory Practices,” *Principles for Responsible Investment*, unpri.org, (2016 & 2021).
12. Ang, A., J. Liu and K. Schwarz, “Using Stocks or Portfolios in Test of Factor Models,” San Francisco Meetings Paper, Jacobs Levy Equity Management Center for Quantitative Financial Research Paper (2018) https://ssrn.com/abstract=1106463
13. Artmann, S., P. Finter, A. Kempf, “Determinants of Expected Stock Returns: Large Sample Evidence from the German Market,” *Journal of Business Finance & Accounting* 39, (5-6) (2012): 758-784.
14. Ashbaugh-Skaife H., D.W. Collins, and R. Lafond, "The Effects of Corporate Governance on Firms' Credit Ratings," *Journal of Accounting & Economics*, volume 42, issue 1-2, (2006): 203 – 243.
15. Attig, N., S. El Ghoul, and O. Guedhami, "Corporate Social Responsibility and Credit Ratings," *Journal of Business Ethics,* Vol 117 (2013): 679-94. https://doi.org/10.1007/s10551-013-1714-2.
16. Auer, B.R, "Do socially responsible investment policies add or destroy European stock portfolio value?" *Journal of Business Ethics* 135(2) (2016): 381–397.
17. Auer, B.R., and F. Schuhmacher, "Do socially (ir) responsible investments pay? New evidence from international ESG data," *The Quarterly Review of Economics and Finance* 59 (2016): 51-62.

18. Aupperle, K.E., A.B. Carroll, and J.D. Hatfield, "An empirical examination of the relationship between corporate social responsibility and profitability," *Academy of Management Journal* 28(2)(1985): 446–463.
19. Bai, J., T.G. Bali, and Q. Wen, "Common risk factors in the cross-section of corporate bond returns," *Journal of Financial Economics,* (2018).
20. Baker, Malcolm, Brendan Bradley, and Jeffrey Wurgler, "Benchmarks as Limits to Arbitrage: Understanding the Low-Volatility Anomaly," *Financial Analysts Journal,* Volume 67, Issue 1, (2011).
21. Baker, Malcolm, Brendan Bradley, and Jeffrey Wurgler, "Benchmarks as Limits to Arbitrage: Understanding the Low-Volatility Anomaly," *Financial Analysts Journal,* Volume 67, Issue 1, (2011).
22. Baker, N.L. and R. A. Haugen A. Robert, "Low Risk Stocks Outperform within All Observable Markets of the World," (2012) https://ssrn.com/abstract=2055431
23. Banz, Rolf W., "The Relationship Between Return and Market Value of Common Stocks," *Journal of Financial Economics*, 9 (1) (1981): 3-18.
24. Bao, J., J. Pan, and J. Wang, "The Illiquidity of Corporate Bonds," *The Journal of Finance,* 66(3) (2011): 911-946. https://doi.org/10.1111.j.1540-6261.2011.01655.x .
25. Barber, J., "Mapping the movement to achieve sustainable production and consumption in North America," *Journal of Cleaner Production* 15(6) (2007):499-512.
26. Barnea, A. and A. Rubin, "Corporate Social Responsibility as a Conflict Between Owners, *Working Paper, University of British Columbia* (2005).
27. Barney, J. B., "Why resource-based theory's model of profit appropriation must incorporate a stakeholder perspective*," Strategic Management Journal* 39(13) (2018):3305-3325.
28. Baron, D.P., "Private Politics, Corporate Social Responsibility, and Integrated Strategy," *Journal of Economics & Management Strategy*. 10(1) (2001): 7-45. Doi:10.1162/105864001300122548.
29. Barth F., and B. Huebel, H. Scholz, "ESG and corporate credit spreads," (SSRN) *Papers.* Posted (2018, 2021).
30. Basu, S., "The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence," *Journal of Financial Economics 12 (1) (1983): 129-156.*
31. Bauer, R. and D. Hann, "Corporate Environmental Management and Credit Risk," Working paper, Maastricht University (2014): 44 pp https://ssrn.com/abstract=1660470. https://doi.org/10.2139/ssrn.1660470.
32. Bauer, R., K. Koedijk, and R. Otten, "International evidence on ethical mutual fund performance and investment style," *Journal of Banking & Finance* 29(7)(2005): 1751–1767.
33. Bebchuk, L. A., A. Cohen, and A. Ferrell, A, "What Matters in Corporate Governance?" *Review of Financial Studies,* 22(2) (2009): 783-827. https://doi.org/10.1093/rfs/hhn099.
34. Bektic, D., J. Wenzler, M. Wegener, D. Schiereck, and T. Spielmann, "Extending Fama-French Factors to Corporate Bond Markets", *working paper*, (2016): ssrn.com/abstract= 2715727.
35. Bektic, D.R., "ESG Factors in Corporate Bond Returns: Perspectives for Academic Research and Investors," *Journal of Environmental Law and Policy* 40(4) (2018): 293-8.
36. Bektic, D.R., "ESG Factors in Corporate Bond Returns: Perspectives for Academic Research and Investors," *Journal of Environmental Law and Policy*, 40 (4) (2018): 293 – 301.
37. Benabou,R. and J. Tirole, "Individual and corporate social responsibility," *Economica*, (2010).
38. Benartzi, S., R. Michaely, and R. Thaler, "Do changes in dividends signal the future or the past?" *The Journal of Finance* 52(3) (1997): 1007-1034.
39. Benlemlih,M., "Why do socially responsible firms pay more dividends," *PWC*, (2014).
40. Bennani, L., T. Le Guenedal, F. Lepetit, L. Ly, V. Mortier, T. Roncalli, and T. Sekine, "How ESG Investing Has Impacted the Asset Pricing in the Equity Market," (SSRN*) Papers* (2019).

41. Berens, G., C.B.M. van Riel, and G.H. van Bruggen, “Corporate Associations and Consumer Product Responses: The Moderating Role of Corporate Brand Dominance*,” Journal of Marketing,* 69 (2005): 35-48.
42. Berg, F., J. Kölbel, and R. Rigobon, "Aggregate Confusion: The Divergence of ESG Ratings," (2020): https://ssrn.com/abstract=3438533 or http://dx.doi.org/10.2139/ssrn.3438533.
43. Bessembinder, H., K. M. Kahle,. M., W.F. Maxwell, and D. Xu, "Measuring abnormal bond performance," *Review of Financial Studies* 22, (2009): 4219–4258.
44. Bessembinder, H., W. F. Maxwell, K. Venkataraman," Market transparency, liquidity externalities, and institutional trading costs in corporate bonds," J*ournal of Financial Economics* 82, (2006): 251–288.
45. Bharath, S., Shumway, T., “Forecasting default with the Merton distance to default model,” *Review of Financial Studies* 21 (2008): 13391369.
46. Bhojraj, S., and P. Sengupta, "Effect of Corporate Governance on Bond Ratings and Yields: The Role of Institutional Investors and Outside Directors," *The Journal of Business,* 76(3) (2003): 455-475. www.jstor.org/stable/10.1086/344114.
47. Binsbergen, Jules H. van, and Ralph S.J. Koijen, "The Term Structure of Returns: Facts and Theory," *NBER Working Papers* 21234, National Bureau of Economic Research, Inc. (2015).
48. Binsbergen, Jules H. van, and Ralph S.J. Koijen, "The Term Structure of Returns: Facts and Theory," *NBER Working Papers* 21234, National Bureau of Economic Research, Inc. (2015).
49. Black, F., M.C. Jensen, and M. Scholes, “The Capital Asset Pricing Model: Some Empirical Tests,” In Studies in the Theory of Capital Markets, edited by M.C. Jensen. New York: Praeger, (1972).
50. Blitz, D., J. Pang and P. Van Vliet,” the volatility effect in emerging markets,” Emerging Markets Review 16 (2013): 31-45.
51. Blitz, D., Huji, J., Lansdorp, S., van Vliet, P., Factor Investing: Long-Only versus Long-Short (March 28, 2014). Available at SSRN: https://ssrn.com/abstract=2417221 or http://dx.doi.org/10.2139/ssrn.2417221
52. Boehmer, E., and E.K. Kelley, «Institutional Investors and the Informational Efficiency of Prices,” The Review of Financial Studies, 22(9) (2009): 3563-3594
53. Bollen, N.P.B., “Mutual Fund Attributes and Investor Behavior,” *Journal of Financial and Quantitative Analysis* 42 (03) (2007): 683-708.
54. Borgers, A., J. Derwall, K. Koedijk, and J. Ter Horst, "Stakeholder Relatio ns and Stock Returns: On Errors in Investors' Expectations and Learning," *Journal of Empirical Finance* 22, (2013): 159-175. Doi:10.1016/j.jempfin.2013.04.003.
55. Bouslah,K., L. Kryzanowski, and B. Mzali, “The Impact of the Dimensions of Social Performance on Firm Risk,” *Journal of Banking & Finance* 37 (4) (2011).
56. Brammer, S. and A. Millington, "Firm size, organizational visibility and corporate philanthropy: An empirical analysis," *Business Ethics: A European Review, 15*(1) (2006): 6–18. https://doi.org/10.1111/j.1467-8608.2006.00424.x.
57. Breuer, H., K. Fichter, F.Luedeke-Freund and I.Tiemann, “Sustainability-Oriented Business Model Development: Principles, Criteria, and Tools”*International Journal of Entrepreneurial Venturing* 10 (2) (2018): 256-286.
58. Breusch, T.S, "Testing for autocorrelation in dynamic linear models," *Australian Economic Papers* 17(31) (1978): 334–355.
59. Breusch, T.S., and A.R. Pagan, "A simple test for heteros cedasticity and random coefficient variation," *Econometrica* 47 (1979): 1287–1294.
60. Brown, T.J., and PA Dacin, “The company and the product: Corporate associations and consumer product responses,” *Journal of Marketing* 61 (1) (1997): 68-84.

61. Brown, W., E. Helland, and K. Smith, "Corporate philanthropic practices," Journal of Corporate Finance 12 (2006): 855 – 877.
62. Bucholz, R.A. and G.J. Alexander, "Corporate social responsibility and stock market performance," *Academy of Management journal* 21 (3) (1978): 479-486.
63. Campbell, J.Y., A.W. Lo and A.C. MacKinlay, R.F. Whitelaw, "The econometrics of financial markets," *Macroeconomic Dynamics* 2(4) (1998): 559-562.
64. Campello, M., L.Chen, L.Zhang, "Expected returns, yield spreads, and asset pricing test," *The Review of Financial Studies* 21(3) (2008): 1297-1338.
65. Capelle-Blancard, G. and S. Monjon, "Trends in the literature on socially responsible investment: looking for the keys under the lamppost," *Business Ethics: A European Review* 21 3(2012): 239-250.
66. Carhart, M.M, "On persistence in mutual fund performance," *The Journal of Finance* 52(1) (1997): 57–82.
67. Carvalho et al. (2015) de Carvalho Ferreira, M.C.R., Sobreiro, V.A., Kimura, H., de Moraes Barboza, F.L.: A systematic review of literature about finance and sustainability. J. Sustain. Finance Invest. 6(2), 112–147 (2016)
68. Carvalho, M.C.R. Ferreira de, V.A. Sobreiro, H. Kimura, and F.L. de Moraes Barboza, "A systematic review of literature about finance and sustainability," *J. Sustain. Finance Invest.* 6(2) (2016): 112–147.
69. Charness, G. and M. Rabin, "Understanding Social Preferences with Simple Tests," *The Quarterly Journal of Economics* 117(3) (2002) 817-869.
70. Chatterji, A.K., R. Durand, D.I. Levine, and S. Touboul, "Do ratings of firms converge? Implications for managers, investors and strategy researchers," *Strategic Management Journal* 37(8) (2016): 1597–1614.
71. Chava. S., "Environmental Externalities and Cost of Capital," *Management Science* 60(9) (2014): 2223- 47. https://doi.org/10.1287/mnsc.2013.1863.
72. Chen, L., D. A. Lesmond, and J. Wei, "Corporate Yield Spreads and Bond Liquidity," *The Journal of Finance* 62(1) (2007): 119-149.
73. Chen, N., Roll, R., Ross, S., "Economic forces and the stock markets," *Journal of Business* 59 (1986): 383–403.
74. Chen, N.-F., R. Roll, and S. Ross, Economic Forces and the Stock Market, *The Journal of Business*, Vol. 59, No. 3 (1986): 383-403.
75. Chen, Z. and B. Kim, "Fundamental Risk Sources and Pricing Factors," *Nanyang Technological University ABFER* (2019).
76. Chen, Z., A. A. Lookman, N. Schürhoff, and D. J. Seppi, "Rating-Based Investment Practices and Bond Market Segmentation," *The Review of Asset Pricing Studies* 4(2) (2014): 162-205.
77. Cheng B., I. Ioannou, and G. Serafeim, "Corporate Social Responsibility and Access to Finance," *Strategic Management Journal* 3 (1) (2014, 2020): 1 – 23.
78. Cho,S.Y., C. Lee and R. J. Pfeiffer Jr "Corporate social responsibility performance and information asymmetry," *Journal of Accounting and Public Policy* (2013).
79. Chordia, T., A. Goyal, Y. Nozawa, A. Subrahmanyam, and Q. Tong, "Are capital market anomalies common to equity and corporate bond markets?" *Journal of Financial and Quantitative Analysis* 52 (2017): 1301–1342.
80. Cici, G. and S. Gibson," The Performance of Corporate Bond Mutual Funds: Evidence Based on Security-Level Holdings*," Journal of Financial and Quantitative Analysis* 47(1) (2012): 159-178.
81. Cochrane, J.H., and M. Piazzesi," Bond Risk Premia*," American Economic Review* 95(1) (2005): 138-160.

82. Collin-Dufrense, P. and R.S. Goldstein, “Do credit spreads reflect stationary leverage ratios?” *The journal of finance* 56(5) (2001): 1929-1957.
83. Cooper, I., L. Ma, and P.F. Maio, and P. Dennis” Multifactor Models and Their Consistency with the APT,” *Review of Asset Pricing Studies* (2020) https://ssrn.com/abstract=2784898
84. Cornett, M.M., O. Erhemjamts, and H. Tehranian, "Greed or good deeds: An examination of the relation between corporate social responsibility and the financial performance of U.S. commercial banks around the financial crisis," *Journal of Banking & Finance* 70(C) (2016): 137 – 159.
85. Correia, M., S.Richardson, and I.Tuna, “Value Investing in Credit Markets,” *Review of Accounting Studies* 17(3) (2012): 572–609.
86. Cox, P., S. Brammer and A. Millington, “An Empirical Examination of Institutional Investor Preferences for corporate Social Performance,” Journal of Business Ethics 52 (2004): 27-43.
87. Crabbe L., and C.M. Turner, "Does the Liquidity of a Debt Issue Increase with its Size? Evidence from the Corporate Bond and Medium-Term Note Markets," *The Journal of Finance* 50(5) (1995): 1719 – 1734.
88. Crawford et al., 2015 Crawford, S. S., Roulstone, D. T., & So, E. C. (2012). Analyst initiations of coverage and stock return synchronicity. The Accounting Review,87(5), 1527–1553.
89. Cremars, M. and V. B. Nair, "Governance Mechanisms and Equity Prices," *Journal of Finance* 60(6) (2005): 2859-2894.
90. Cubas-Díaz,M. and M. Ángel Martínez Sedano, "Do Credit Ratings Take into Account the Sustainability Performance of Companies?" *Sustainability* 10(11) (2018): 4272.
91. Cui, B. and P. Docherty, “Stock Price Overreaction to ESG Controversies,” (March 17, 2020). Available at SSRN: https://ssrn.com/abstract=3559915 or http://dx.doi.org/10.2139/ssrn.3559915.
92. De Bondt, W. F. M. and R. Thaler, “Does the Stock Market Overreact? The journal of Finance 40(3) (1985): 793-805.
93. Demers, E., J. Hendrikse, P. Joos and B.I. Lev, "ESG Didn’t Immunize Stocks During the COVID-19 Crisis, But Investments in Intangible Assets Did," *Journal of Business Finance & Accounting* 48 (2021): 433-462. https://ssrn.com/abstract=3675920.
94. Derwall, J. and K. C. G. Koedijk, "Socially Responsible Fixed-Income Funds," *Journal of Business Finance and Accounting* 36 (2009): 210-29. https://doi.org/10.1111/j.1468-5957.2008.02119.x.
95. Derwall, J., N. Guenster, R. Bauer, and K. Koedijk, "The eco-efficiency premium puzzle," *Financial Analysts Journal* 61(2) (2005): 51-63. https://doi.org/10.2469/faj.v61.n2.2716.
96. Dichev, I.D., “Is the Risk of Bankruptcy a Systematic Risk?” *The Journal of Finance* 53(3) (2002): 1131-1147.
97. Dorfleitner, G., C. Kreuzer, and C. Sparrer, "ESG controversies and controversial ESG: about silent saints and small sinners," *J Asset Manag* 21 (2020): 393–412. https://doi.org/10.1057/s41260-020-00178-x.
98. Dorfleitner, G., G. Halbritter, and M. Nguyen, "Measuring the level and risk of corporate responsibility—An empirical comparison of different ESG rating approaches," *Journal of Asset Management* 16(7) (2015): 450–466.
99. Döttling, R. and S. Kim, "Sustainability Preferences Under Stress: Evidence from Mutual Fund Flows during COVID-19," (2021). https://ssrn.com/abstract=3656756.
100. Doukas, John A., et al. “A Test of the Errors-in-Expectations Explanation of the Value/Glamour Stock Returns Performance: Evidence from Analysts' Forecasts,” *The Journal of Finance* 57(5) (2002): 2143–2165. JSTOR, www.jstor.org/stable/3094507. Accessed 3 May 2021.
101. Drempetic, S., C. Klein, and B. Zwergel, "The Influence of Firm Size on the ESG Score: Corporate Sustainability Ratings Under Review," *J Bus Ethics* 167 (2020): 333–360. https://doi.org/10.1007/s10551-019-04164-1.

102. Drut, B., "Sovereign Bonds and Socially Responsible Investment," Journal of Business Ethics 92(1) (2010): 131 – 145.
103. Durbin, J., and G.S. Watson, "Testing for serial correlation in least squares regression," *III. Biometrika* 58(1) (1971): 1–19.
104. Dyck, A. and H.F. Wagner, "Do institutional investors drive corporate social responsibility?" International evidence, Journal of Financial Economics 131(3) (2019): 693-714.
105. Dynkin, L., J. Hyman, V. Konstantinovsky, " Suficient diversification in credit portfolios: Number of issues and downgrade risk," *ResearchGate.net* (2002)
106. Eccles, R. G., I. Ioannou, and G. Serafeim, "The Impact of Corporate Sustainability on Organizational Processes and Performance," *Management Science* 60(11) (2014): 2835-2857. https://ssrn.com/abstract=1964011 or http://dx.doi.org/10.2139/ssrn.1964011.
107. Eding, E., and B. Scholtens, "Corporate Social Responsibility and Shareholder Proposals: Shareholder Proposals," Corporate Social Responsibility and Environmental Management 24(1) (2017): DOI:10.1002/csr.1434.
108. Eding, E., and B. Scholtens, "Corporate Social Responsibility and Shareholder Proposals: Shareholder Proposals," Corporate Social Responsibility and Environmental Management 24(1) (2017): DOI:10.1002/csr.1434.
109. Edmans, A., "Does the stock market fully value intangibles? Employee satisfaction and equity prices," *Journal of Financial Economics* 101(3) (2011): 621–640.
110. El Ghoul, S., and A. Karoui, "Does corporate social responsibility affect mutual fund performance and flows?" *Journal of Banking & Finance 77* (2017): 53-63. https://doi.org/10.1016/j.jbankfin.2016.10.009.
111. Elton, E. J., Gruber, M. J., Das, S., Hlavka, M., "Efficiency with Costly Information: A Reinterpretation of Evidence from Managed Portfolios," *Review of Financial Studies* 6(1) (1993): 1- 22.
112. Elton, E.J., M. J. Gruber, D. Agrawal, and C. Mann, "Explaining the Rate Spread on Corporate Bonds," *The Journal of Finance* 56(1) (2002: 247-277.
113. Eurosif. (2018). *European SRI Study 2018*. Retrieved January 04, 2019 from http://www.eurosif.org/wp-content/uploads/2018/11/European-SRI-2018-Study-LR.pdf. Accessed 4 January 2021.
114. Fama and MacBeth, 1976). Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. Journal of Political Economy,81(3), 607–636.
115. Fama, E., and K. French, "A five-factor asset pricing model," *Journal of Financial Economics* 116(1) (2015): 1–22.
116. Fama, E.F., and K. R.French " The Cross-Section of Expected Stock Returns," *The Journal of Finance* 47(2) (1992, 1993)
117. Fehr, E. and KM Schimdt, "A theory of fairness, competition, and cooperation*," The quarterly journal of economics* 114 (3) (1999): 817-868.
118. Feldhütter, P. and S. M. Schaefer, The «Myth of the Credit Spread Puzzle,» (2018). https://ssrn.com/abstract=2363081
119. Feng, G., S. Giglio, and D. Xiu, "Taming the Factor Zoo: A Test of New Factors," *The Journal of Finance* 75 (3) (2020): 1327-1370.
120. Frazzini, A., and L.H. Pedersen, "Betting against beta," *Journal of Financial Economics* 111(1) (2014.): 1–25.
121. Freeman, E., "Strategic management: A stakeholder approach," Marshfield, MA: Pitman. (1984)
122. Friede, G., T. Busch, and A. Bassen, "ESG and financial performance: Aggregated evidence from more than 2000 empirical studies," *Journal of Sustainable Finance & Investment* 5(4) (2015): 210–233.

123. Friedman, M., "The social responsibility of business is to increase its profits," *The New York Times Magazine* (September 13, 1970): 122-148.
124. Ftse, R., “Smart Sustainability: 2020 Global Survey Findings from Asset Owners,” (2020). Available online: https://www.ftserussell.com/index/spotlight/smart-sustainability-survey (accessed on 16 April 2021).
125. Gao, H., K. Li, and Y. Ma, "Stakeholder orientation and the cost of debt: Evidence from state level adoption of constituency statutes," (2014, 2020).
126. Gavana, G., P. Gottardo, and A. M. Moisello, "The effect of equity and bond issues on sustainability disclosure. Family vs non-family Italian firms," Social Responsibility Journal, (2017): ISSN: 1747-1117.
127. Ge, W. and M. Liu, "Corporate social responsibility and the cost of corporate bonds," *Journal of Accounting and Public Policy* 34(6) (2015)*:* 597-624. https://doi.org/10.1016/j.jaccpubpol.2015.05.008.
128. Gebhardt, W. R., Hvidkjaer, S., Swaminathan, B., “Stock and bond market interaction: Does momentum spill over?” *Journal of Financial Economics* 75 (2005): 651–690.
129. Gerakos, J. and J. T. Linnainmaa, “Decomposing value,”(2016) Fama-Miller Working Paper, Chicago Booth Research Paper No 12-18, Review of Financial Studies 31, no.5 (2018): 1825-1854.
130. Gerard, B., "ESG and Socially Responsible Investment: A Critical Review," (December 28, 2018). Available at https://ssrn.com/abstract=3309650.
131. Gillan, S. L., A. Koch, L.T. Starks, «Firms and Social Responsibility: A Review of ESG and CSR. Research in Corporate Finance, CSB Working Papers Series No. 2020-02.
132. Godfrey, P., C. B Merrill and J.M Hansen, "The relationship between corporate social responsibility and shareholder value: An empirical test of the risk management hypothesis," *Strategic management journal* 30 (4) (2009): 425-445.
133. Goetzmann, W.N. and M .Schaefer, “ Evaluation of Active Management of Norwegian Government Pension Fund-Global (2009.)
134. Gold, S. and P.P.Heikkurinen, “ Transparency Fallacy: Unintended Consequences of Stakeholder Claims on Responsibility in Supply Chains*, Accounting Auditing and Accountability Journal* 31(1) (2018) 318-337. ISSN 0951-3574 https://doi.org/10.1108/AAAJ-06-2015-2088.
135. Goldstein, I., H. Jiang, and D. Ng, “Investor Flows and Fragility in Corporate Bond Funds,” *Journal of Financial Economics* (2017).
136. Gompers, P., J. Ishii, and A. Metrick, "Corporate governance and equity prices," *Quarterly Journal of Economics,* 118(1) (2003): 107–155.
137. Goss, A. and G.S. Roberts, “The impact of corporate social responsibility on the cost of bank loans,” *Journal of Banking & Finance* 35(7) (2011): 1794-1810.
138. Gourio, F., “Macroeconomics”, *American Economic Journal*, 5(3) (2013).
139. Gray, W. B., and R. J. Shadbegian, "Environmental Regulation and Manufacturing Productivity at the Plant Level," *NBER Working Paper* No. w4321. Posted (2001) Written (1993).
140. Greening, D. W. and D. B. Turban, "Corporate Social Performance AS A Competitive Advantage In Attracting A Quality Workforce," *Business & Society* 39(3) (2000): 254-280.
141. Gregory A., R. Tharyan and J. Whittaker, "Corporate Social Responsibility and Firm Value: Disaggregating the Effects on Cash Flow, Risk and Growth," *Journal of Business Ethics* 124(4) (2013): DOI:10.1007/s10551-013-1898-5.
142. Guenster, N., R. Bauer, J. Derwall, and K. Koedijk, "The economic value of corporate eco-efficiency," *European Financial Management* 17(4) (2011): 679–704.
143. Gutierrez, J.A., V. Martinez, Y. Tse, « Where does return and volatility come from? The case of Asian ETFs,” International Review of Economics & finance 18(4) (2009): 671-679.

144. Haesen, D., P. Houweling, and J. van Zundert, “Momentum Spillover from Stocks to Corporate Bonds,” *Working paper*, (2013, 2015): ssrn.com/abstract=2131032.
145. Halbritter, G., and G. Dorfleitner, "The wages of social responsibility—Where are they? A critical review of ESG investing," *Review of Financial Economics* 26 (2015): 25–35.
146. Halling, M., J. Yu, and J. Zechner, "How did covid-19 affect firms' access to public capital markets?" *Review of Corporate Finance Studies 9 (3)* (2020): 501-533.
147. Hart, S.L., “A Natural-Resource-Based View of the Firm*,” The Academy of Management Review* 20(4) (1995): 986-1014.
148. Harvey, C. R., Y. Liu, and H. Zhu,”…and the Cross-Section of Expected Returns,” *The Review of Financial Studies*, 29(1) (2016): 5-68.
149. Haugen, Robert A. and A. James Heins, "Risk and the Rate of Return on Financial Assets: Some Old Wine in New Bottles," *Journal of Financial and Quantitative Analysis*, vol. 10, issue 5 (1975):, 775-784.
150. Heinkel et al. (2001) Heinkel, R., Kraus, A., Zechner, J.: The effect of green investment on corporate behavior. J. Financ. Quant. Anal. 36(4), 431–449 (2001)
151. Henke, H.-M., "The Effect of Social Screening on Bond Mutual Fund Performance," *Journal of Banking and Finance* 67 (2016): 69-84. https://doi.org/10.1016/j.jbankfin.2016.01.010.
152. Hillman, A. J. and G.D. Keim, "Shareholder value, stakeholder management, and social issues: what’s the bottom line?" *Strategic Management Journal* 22(2) (2001): 125-39. https://doi.org/10.1002/1097-0266(200101)22:2< 125::aid-smj150>3.0.co;2-h.
153. Hillman, A. J. and G.D. Keim, "Shareholder value, stakeholder management, and social issues: what’s the bottom line?" *Strategic Management Journal* 22(2) (2001): 125-39. https://doi.org/10.1002/1097-0266(200101)22:2< 125::aid-smj150>3.0.co;2-h.
154. Hirshleifer, D. A., "Psychological Bias as a Driver of Financial Regulation," *European Financial Management Journal*, Posted (2007).
155. Hoepner, A. G. F., D. McMillan, and M. Fraser, "Is responsible investment proportionally underresearched?" In T. Hebb, J. P. Hawley, & A. G. F. Hoepner (Eds.), *The Routledge handbook of responsible investment* (2016): 34–50. (Routledge companions in business, management and accounting).
156. Hoepner, A.G.F. & M. Nilsson, "No news is good news: Corporate social responsibility ratings and fixed income portfolios," *Working Paper, University of Reading*. (2020): https://doi.org/10.2139/ssrn.2943583 .
157. Hoepner, A.G.F. and M. Nilsson, "Expertise among SRI fixed income funds and their management companies," *Working Paper, University of Reading*. (2017): https://doi.org/10.2139/ssrn.2517057 .
158. Hoepner, A.G.F., I. Oikonomou, B. Scholtens, and M. Schröder, "The Effects of Corporate and Country Sustainability Characteristics on the Cost of Debt: An International Investigation," *Journal of Business Finance & Accounting* 43 (2016): 158-90. https://doi.org/10.1111/jbfa.12183 .
159. Hoepner, A.G.F., I. Oikonomou, Z. Sautner, L.T. Starks, and X. Zhou, "ESG Shareholder Engagement and Downside Risk," *AFA paper* (2018): https://doi.org/10.2139/ssrn.2874252.
160. Holt, D. and R. Barkmeyer, L. Preuss, and S. Tsang, " What happened to the ‘development’ in sustainable development? Business guidelines two decades after Brundtland," *Sustainable Development* 22(1) (2014): 15-32.
161. Hong, H. G. and Sraer, A., Speculative Betas (November 2012). NBER Working Paper No. w18548, Available at SSRN: https://ssrn.com/abstract=2178327
162. Hong, H. G., and I. Liskovich, "Crime, Punishment and the Value of Corporate Social Responsibility," *Working paper*, Columbia University (2016). https://doi.org/10.2139/ssrn.2492202.

163. Hong, H., and M. Kacperczyk, "The price of sin: The effects of social norms on market," *Journal of Financial Economics* 93(1) (2009): 15–36.
164. Hong, H., Stein, J. C., "A unified theory of underreaction, momentum trading, and overreaction in asset markets," *Journal of Finance* 54 (1999): 2143–2184.
165. Horowitz, J. L., Loughran, T., & Savin, N. E. (2000). The disappearing size effect. Research in Economics, 54(1), 91-116. https://doi.org/10.1006/reec. 1999.0207
166. Hörter, S., "ESG in Investment Grade Corporate Bonds," *Working Paper, Allianz Global Investors GmbH*. (2016,2017).
167. Houweling, P. and J. van Zundert, "Factor Investing in the Corporate Bond Market (September 26, 2016). Financial Analysts Journal, 2017, Vol. 73, No. 2, Available at SSRN: https://ssrn.com/abstract=2516322 or http://dx.doi.org/10.2139/ssrn.2516322
168. Huang, J.-Z., Huang, M. How much of the corporate-treasury yield spread is due to credit risk? (2003) Unpublished working paper. Pennsylvania State University.
169. Hübel, B. and H.Scholz, "Integrating sustainability risks in asset management: the role of ESG exposures and ESG ratings," *Journal of Asset Management* 21 (1) No 5, (2020): 52-69.
170. Huberman Gur & Zhenyu Wang, "Arbitrage pricing theory," *Staff Reports* 216, Federal Reserve Bank of New York. (2005).
171. Huij J. and J. Derwall, "Hot Hands" in bond funds," *Journal of Banking & Finance* 32 (4)(2008)): 559-572.
172. Iaonnou I. and G. Serafeim, "What drives corporate social performance? The role of national-level institutions," *Journal of International Business Studies* 43, (2012): 1-31.
173. Iaonnou I. and G. Serafeim, "What drives corporate social performance? The role of national-level institutions," *Journal of International Business Studies* 43, (2012): 1-31.
174. Ielasi, F., P. Ceccherini and P. Zito, "Integrating ESG Analysis into Smart Beta Strategies," *Sustainability* 12(22) (2020): 9351 https://doi.org/10.3390/su12229351.
175. Ilmanen Antti, Rory Byrne, Heinz Gunasekera, and Robert Minikin, "Which Risks have Been Best Rewarded?" *The Journal of Portfolio Management* 30(2)(2004): 53-57. DOI:10.3905/jpm.2004.319930
176. Ilmanen Antti, Rory Byrne, Heinz Gunasekera, and Robert Minikin, "Which Risks have Been Best Rewarded?" *The Journal of Portfolio Management* 30(2)(2004): 53-57. DOI:10.3905/jpm.2004.319930.
177. IMF.org, Financial Stability Report, (2019).
178. Israel, R., D. Palhares, and S. Richardson, "Common factors in corporate bond returns," *The Journal of Investment Management.* 51 Pages, Last revised: 31 Aug 2017 (2018).
179. Jagannathan R., A Ravikumar, M Sammon, "Environmental, Social, and Governance Criteria: Why Investors are Paying Attention," *NBER Working Paper* No. w24063 (2017).
180. Jayachandran, S., K. Kalaignanam, and M. Eilert, "Product and environmental social performance: Varying effect on firm performance," *Strategic Management Journal* 34(10) (2013): 1255-1264. https://doi.org/10.1002/smj.2054
181. Jeffers, A.E., L. Beixin, S. Romero, and L. A DeGaetano, "Sustainability (green) initiatives and company's value-an examination of S&P 500 US companies," *International Journal of Society systems science* 7(3) (2015): 181-197.
182. Jegadeesh, N., Titman, S., "Returns to buying winners and selling losers: implications for stock market efficiency", *Journal of Finance* 48(1) (1993): 65–91.
183. Jensen M., and W. Meckling, "Theory of the firm: Managerial behavior, agency costs and ownership structure," *Journal of Financial Economics* 3(4) (1976): 305 – 360.
184. Jiao, Y., "Stakeholder welfare and firm value," Journal of Banking & Finance 34(10) (2010): 2549-2561.

185. Jin, Ick, “Is ESG a systematic risk factor for US equity mutual funds?” *Journal of Sustainable Finance & Investment* 8 (2018): 72-93

186. Jiraporn, P., N. Jiraporn, A. Boeprasert, K. Chang, “Does Corporate Social Responsibility (CSR) Improve Credit Ratings? Evidence from Geographic Identification,” *Financial Management* 43(3) (2013): 505-531.

187. Jones, T.M., “Instrumental Stakeholder Theory: A Synthesis of Ethics and Economics,” *The Academy of Management Review*, 20(2) (1995): 404-437. https://doi.org/10.2307/258852.

188. Jostova, G., S. Nikolova, A. Philipov, and C. Stahel, “Momentum in corporate bond returns,” Review of Financial Studies 26 (2013): 1649–1693.

189. Kapadia, Nikunj, and Xiaoling Pu, "Limited arbitrage between equity and credit markets," *Journal of Financial Economics,* (2012): 542-564.

190. Kempf, A., and P. Osthoff, "The effect of socially responsible investing on portfolio performance," *European Financial Management,* 13(5) (2007): 908-922.

191. Khan, M., G. Serafeim, and A. Yoon, "Corporate Sustainability: First Evidence on Materiality," *The Accounting Review* 91(6) (2016): 1697-1724. doi: 10.2308/

192. Kim, Y., and M. Statman, "Do corporations invest enough in environmental responsibility?" *Journal of Business Ethics,* 105 (2012): 115-129.

193. King, A. A., and M. J. Lenox, “Exploring the Locus of Profitable Pollution Reduction,” *Management Science* 48(2) (2002): 289-299.

194. Klein, J. and N. Dawar, " Corporate Social Responsibility and Consumers' Attributions and Brand Evaluations in a Product-Harm Crisis," *International Journal of Research in Marketing* 21 (2004): 203-217. http://dx.doi.org/10.1016/j.ijresmar.2003.12.003.

195. Klock, M., S. Mansi, and W. Maxwell, “Does corporate governance matter to bondholders?” *Journal of Financial and Quantitative Analysis* 40(4) (2005): 693 – 719.

196. Koh, S., M. Limkriangkrai, and R. B. Durand, “Environmental, Social, and Governance (ESG) Profiles, Stock Returns, and Financial Policy: Australian Evidence,” *International Review of Finance,* 17(3) (2016): 461-471.

197. Konar, S. and M.A. Cohen, "Does the market value environmental performance?" *Review of Economics and Statistics*, 83(2) (2001): 281-89. https://doi.org/10.1162/00346530151143815.

198. Kozhemiakin, A. “The Risk Premium of Corporate Bonds”, The Journal of Portfolio Management Winter 2007, 33 (2) 101-109; DOI: https://doi.org/10.3905/jpm.2007.674796

199. Krüger, P., "Corporate goodness and shareholder wealth," *Journal of Financial Economics,* 115(2) (2015): 304–329.

200. Kumbhat, A., F. Palomino, and Ander Perez-Orive, "The Potential Increase in Corporate Debt Interest Rate Payments from Changes in the Federal Funds Rate," *FEDS Notes.* Washington: Board of Governors of the Federal Reserve System, November 15, 2017: https://doi.org/10.17016/2380-7172.2076.

201. La Torre, M., S. Sabelfeld, M. Blomkvist and J. Dumay, "Rebuilding trust: sustainability and non-financial reporting and the European Union regulation," *Meditari Accountancy Research* 28(5) (2020) ISSN: 2049-372X.

202. Lakonishok, J. A. Shleifer, R. Vishny, *“*Contrarian Investment, extrapolation, and Risk*”*, *The Journal of Finance* 49(5) (1994): 1541-1578

203. Lakonishok, J., A. Shleifer, and R. Vishny, "Contrarian investment, extrapolation, and risk,"

204. Lee, D., Faff, R., Rekker, S.: Do high and low-ranked sustainability stocks perform differently? Int. J. Account. Inf. Manag. 21(2), 116–132 (2013)

205. Lee, D.D. and R.W. Faff, “Corporate Sustainability Performance and Idiosyncratic Risk: A Global Perspective”, *Financial Review* 44, (2009): 213-237

206. Leite, P., and M.C. Cortez, "The Performance of European Socially Responsible Fixed Income Funds," *Investment Analysts Journal*, 47(1) (2018): 65-79. https://doi.org/10.1080/10293523.2017.1414911.
207. Leland, H.E, and K.B. Toft, "Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads", *The Journal of Finance* 51, (1996): 987-1019
208. Lewellen, J., S. Nagel, and J. Shanken, "A skeptical appraisal of asset-pricing tests," *Journal of Financial Economics, 96* (2010): 175-194.
209. Li, G., Zhang, M., Koh, L., The impact of environmental, social, and governance disclosure on firm value: The role of CEO power. (2017)The British Accounting Review. 50. 10.1016/j.bar.2017.09.007.
210. Lins, K. V., H. Servaes, A. Tamayo, "Social Capital, Trust, and Firm Performance: The Value of Corporate Social Responsibility during the Financial Crisis," *The Journal of Finance* 72, (2017): 1785 – 1824.
211. Lintner, J., "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *The Review of Economics and Statistics* 47(1) (1965): 13-37.
212. Lioui, A., "The decarbonisation factor: a new academic fiction?", *EDHEC Working Paper* (2019)
213. Lioui, A., and P. Patrice, "Optimal benchmarking for active portfolio managers", *European Journal of Operational Research* 226 (2), (2013): 268-276.
214. Lipton Amy F. Lipton and Richard J. Kish, "Robust performance measures for high yield bond funds," *The Quarterly Review of Economics and Finance*, vol. 50, issue 3, (2010): 332-340. http://www.sciencedirect.com/science/article/pii/S1062-9769(10)00021-9.
215. Lipton, A., and R. Kish, "Robust performance measures for high yield bond funds", *The Quarterly Review of Economics and Finance* 50(3), (2010): 332-340
216. Litterman Robert B., and Joé Scheinkman, "Common Factors Affecting Bond Returns," *The Journal of Fixed Income* 1 (1)(1991): 54-61. DOI: https://doi.org/10.3905/jfi.1991.692347.
217. Longstaff, F. A., Mithal, S., Neis, E., "Corporate yield spreads: default risk or liquidity? new evidence from the credit-default-swap market," *Journal of Finance* 60 (2005): 2213–2253.
218. Magnanelli, B.S., and M. F. Izzo, "Corporate social performance and cost of dept: the relationship," *Social Responsibility Journal* 13(2) (2017): 250-265.
219. Malkiel, B., "Socially Responsible Investing," speech to 1971 Endowment Conference, reprinted in Ellis ed., Classics II: Another Investor's Anthology, Homewood, Illinois, *AIMR/ Business One Irwin* (1991).
220. Margolis, J. and J. Walsh. "Misery Loves Companies: Rethinking Social Initiatives by Business." Administrative Science Quarterly 48 (2) (2003): 268-305. Doi: 10.2307/3556659.
221. Martenelli, Alberto and Atle Midttun, "Globalization and governance for sustainability," *Corporate Governance International Journal of Business in Society* 10(1)(2010): 6-17. DOI:10.1108/14720701011021076.
222. Masulis, R. and S. Reza, "Private benefits and corporate investment and financing decisions: The case of corporate philanthropy," *European Corporate Governance Institute (ECGI) -Finance Working Paper* (2019).
223. Masulis, R., and S. Reza, "Agency problems of corporate philanthropy," *The Review of Financial Studies* 28(2) (2015): 592 – 636.
224. Matallin-Saez, J., Soler-Dominguez, A., and E. Tortosa-Ausina, "Ethical strategy focus and mutual fund management: performance and persistence," *Working Papers* 1, (2016).
225. McWilliams, A., and D. Siegel, "Corporate social responsibility and financial performance: Correlation or misspecification," *Strategic Management Journal,* 21(5) (2000): 603-609.
226. Mena, S. and G. Palazzo, "Input and Output Legitimacy of multi-Stakeholder Initiatives," *Business Ethics Quarterly* 22(3) (2012): 527-556.

227. Menz, K., "Corporate social responsibility: is it rewarded by the corporate bond market? A critical note," *Journal of Business Ethics* 96 (2010): 117-34. https://doi.org/10.1007/s10551-010-0452-y.
228. Merton, R., "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *The Journal of Finance* 29(2) (1974): 449 – 470.
229. Milevsky, M., A. Aziz, A. Goss, J. Comeault, and D. Wheeler, “Cleaning a Passive Index: How to Use Portfolio Optimization to Satisfy CSR Constraints,” *Journal of Portfolio Management* 32 (3) (2006): 110-120.
230. Mossin, J., “Equilibrium in a Capital Asset Market,” *Econometrica* 34(4) (1966): 768-783.
231. Nagy, Z., A. Kassam, and L.E. Lee, “Can ESG Add Alpha? An Analysis of ESG Tilt and Momentum Strategies,” *The Journal of Investing*, 25(2) (2016): 113-124.
232. Nauman, B., "Credit rating agencies join battle for ESG supremacy," https://www.ft.com/content/59f60306-d671-11e9-8367-807ebd53ab77. Accessed: 24th January 2020.
233. Newey, W.K., and K.D. West, "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica* 55(3) (1987): 703-708.
234. Nofsinger, J., and A. Varma, "Socially responsible fund and market crisis," *Journal of Banking & Finance* 48 (C) (2014): 180 – 193.
235. Oikonomou, I., C. Brooks, and S. Pavelin, "The Effects of Corporate Social Performance on the Cost of Corporate Debt and Credit Ratings," *The Financial Review* 49 (2014): 49-75. https://doi.org/10.1111/fire.12025.
236. Orlitzky, M., "Does firm size confound the relationship between corporate social performance and firm financial performance?" *Journal of Business Ethics,* 33 (2001): 167-180.
237. Pagano,M. and P.F. Volpin, "Managers, workers, and corporate control," *The journal of finance* 60 (2) (2005): 841-868.
238. Pastor L, R. Stambaugh, and L. Taylor.,"Sustainable investing in equilibrium," *NBER Working Paper Series* 26549 (2020): http://www.nber.org/papers/w26549 .
239. Patten, D., " The Relation Between Environmental Performance and Environmental Disclosure: A Research Note,” *Accounting Organizations and Society* 27(8) (2002): 763-773.
240. Patton, A., Politis, D., White, H.: Correction to automatic block-length selection for the dependent bootstrap by D. Politis and H. White. Econ. Rev. 28, 372–375 (2009)
241. Penman, S. H., Richardson, S. A., Tuna, A. I., The Book-to-Price Effect in Stock Returns: Accounting for Leverage. Journal of Accounting Research, Vol. 45, No. 2, pp. 427-467, May 2007, Available at SSRN: https://ssrn.com/abstract=1065873 or http://dx.doi.org/10.1111/j.1475-679X.2007.00240.x
242. Pereira, P., M.C. Cortez and F. Silva, "Socially responsible investing and the performance of Eurozone corporate bond portfolios," *Corporate Social Responsibility and Environmental management* 26(6) (2019): 1407-1422.
243. Phillips, C.B., and D. Sul, "Transition Modeling and Econometric Convergence Tests," *Econometrica* 75(6) (2007): 1771-1855. https://doi.org/10.1111/j.1468-0262.2007.00811.x
244. Porter, M.E. and C. Van der Linde, "Toward a new conception of environment-competitiveness relationship," Journal of economic perspectives 9(4) (1995): 97-118.
245. Pospisil, L. and J. Zhang, "Momentum and Reversal Effects in Corporate Bond Prices and Credit Cycles," *Journal of Fixed Income,* 20 (2010): 101-115.
246. Qian, W., and S. Schaltegger, "Revisiting carbon disclosure and performance: Legitimacy and management views," *The British Accounting Review* 49(4) (2017): 365-379. https://doi.org/10.1016/j.bar.2017.05.005.
247. Qiu, Y., A. Shaukat, and R. Tharyan, "Environmental and social disclosures: Link with corporate financial performance," *The British Accounting Review* 48(1) (2016): 102-116.

248. Renneboog, L., J.R. Ter Horst, and C.C. Zhang, "Socially responsible investments: Institutional aspects, performance, and investor behavior," *Journal of Banking & Finance* 32 (2008): 1723-1742.
249. Revelli, C., "Socially responsible investing (SRI): From mainstream to margin?" *Research in International Business and Finance* 39(2017): 711-717. https://doi.org/10.1016/j.ribaf.2015.11.003.
250. Rodriguez-Fernandez, M., "Social responsibility and financial performance: The role of good corporate governance," *BRQ Business Research Quarterly* 9(2) (2016): 137-151. https://doi.org/10.1016/j.brq.2015.08.001.
251. Roll, Richard, "A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory," *Journal of Financial Economics*, 4 (2) (1977): 129–176. doi:10.1016/0304-405X(77)90009-5.
252. Rosenberg, B., "Extra-Market Components of Covariance in Security Returns," *the Journal of Financial and Quantitative Analysis* 9(2) (1974): 263-274.
253. Ross, S., "Risk, return and arbitrage. Risk Return in Finance," ed*. I. Friend and J. Bicksler,* Cambridge, Mass.: Ballinger (1976).
254. Ross, S., "The arbitrage theory of capital asset pricing," *Journal of Economic Theory* 13 (1976):341–60.
255. Rudd, A., "Social Responsibility and Portfolio Performance," *California Management Review*. (1981) https://doi.org/10.2307/41164931
256. Salvi, A., A. Giakoumelou, and G.S.Bertinetti, "CSR in the bond market: Pricing stakeholders and the moderating role of the institutional context*," Global Finance Journal*, 100522 (2020)
257. Sandulescu, M. "How Integrated Are Corporate Bond and Stock Markets?" Ross School of Business Paper Forthcoming Swiss Finance Institute Research Paper No. 20-09, Posted (2020
258. Sandulescu, M., "How Integrated Are Corporate Bond and Stock Markets?" *Ross School of Business Paper Forthcoming Swiss Finance Institute Research Paper No. 20-09*, Posted (2020).
259. Sarig, O. and A. Warga, "Bond Price Data and Bond Market Liquidity," The Journal of Financial and Quantitative Analysis 24(3) (1989): 367-378 https://doi.org/10.2307/2330817.
260. Schaefer Stephen M. and Ilya Strebulaev, "Structural models of credit risk are useful: Evidence from hedge ratios on corporate bonds*," Journal of Financial Economics*, vol. 90, issue 1, (2008): 1-19. http://www.sciencedirect.com/science/article/pii/S0304-405X(08)00100-1.
261. Schäfer, H., A. Hauser-Ditz, and E.-C. Preller, "Transparenzstudie: zur Beschreibung ausgewählter international verbreiteter Rating-Systeme zur Erfassung von Corporate Social Responsibility," (2004). Retrieved February 21, 2021. Accessed 04 January 2021.
262. Schwert,W.G., "Anomalies and Market Efficiency," in Handbook of the Economics of Finance. Eds. G.M. Constantinides, M. Harris, R. Schulz. (2003): 937-972.
263. Seltzer L., L. Starks, and Q. Zhu, "Climate regulatory risks and corporate bonds," *Working paper, University of Texas at Austin*. (2019).
264. Sen, S. & C.B. Bhattacharya, "Does Doing Good Always Lead to Doing Better? Consumer Reactions to Corporate Social Responsibility," *Journal of Marketing Research*, (2001) https://doi.org/10.1509/jmkr.38.2.225.18838
265. Sharfman, M. P., & C.S. Fernando, "Environmental risk management and the cost of capital," *Strategic Management Journal* 29 (2008): 569-592.
266. Sharfman, M.P., and C.S.Fernando, "Environmental risk management and the cost of capital," *Strategic management journal*, (2008) Wiley Online Library
267. Sharpe, L. J., "Theories and Values of Local Government", (1970).
268. Sharpe, W.F.," Risk-Aversion in the Stock Market: Some Empirical Evidence," *The Journal of Finance* 20 (3) (1965): 416-422.

269. Shih, Y., Chen, S., Lee, C., Chen, P., „The Evolution of Capital Asset Pricing Models“, Review of Quantitative Finance and Accounting 42 (2015) 415-448
270. Siegel, D. S., D. Waldman, and A. Link, “Assessing the impact of organizational practices on the relative productivity of university technology transfer offices: an exploratory study,” Research Policy 32(1) (2003): 27-48.
271. Slimane, B., M. de Jong, J.M. Dumas, H. Fredj, T. Sekine, and M. Srb, "Traditional and Alternative Factors in Investment Grade Bonds," *Working Paper*, (2018).
272. Slimane, B., Guenedal, T., Roncalli, T., Thierry, Takaya, S., ESG Investing in Corporate Bonds: Mind the Gap (November 30, 2019). Available at SSRN: https://ssrn.com/abstract=3683472 or http://dx.doi.org/10.2139/ssrn.3683472
273. Blyth, S., Szigety, M. C., & Xia, J. (2016). Flexible indeterminate factor-based asset allocation. Journal of Portfolio Management, 42, 79–93.
274. Starks L.T.P., P. Venkat, and Q. Zhu, "Corporate ESG profiles and investor horizons." Unpublished working paper, (2018).
275. Statman and Glushkov, “The wages of social responsibility,” *Financial Analysts Journal* 65 (4) (2009): 33 – 46.
276. Stein, J., Takeover Threats and Managerial Myopia, Journal of Political Economy, 96, issue 1, p. 61-80, (1988) https://EconPapers.repec.org/RePEc:ucp:jpolec:v:96:y:1988:i:1:p:61-80.
277. Stellner, C., C. Klein, C., & B. Zwergel, "Corporate social responsibility and Eurozone corporate bonds: The moderating role of country sustainability," *Journal of Banking & Finance,* 59 (2015): 538-549. https://doi.org/10.1016/j.jbankfin.2015.04.032.
278. Sun, W. and K. Cui, “Linking corporate social responsibility to firm default risk,” *European Management Journal*, Elsevier, 32(2) (2014): 275-287.
279. Tang D.Y., and H. Yan, "Market conditions, default risk and credit spreads." *Journal of Banking & Finance* 34, (2010): 743-753.
280. Thomson Reuters, "Environmental, social and governance (ESG) scores from Refinitiv," (2019) https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-scores-methodology.pdf. Accessed 25 Oct 2019.
281. Treynor, J.L., “Market Value, Time and Risk,” (1961). Available at SSRN: http://ssrn.com/abstract=2600366
282. Turban, D. B., and D.W. Greening, “Corporate social performance and organizational attractiveness to prospective employees,” A*cademy of Management Journal* 40, (1997): 658-672.
283. Ullman, A. A., "Data in search of a theory: A critical examination of the relationships among social performance, social disclosure, and economic performance," *Academy of Management Review* 10 (1985): 540-577.
284. UNPRI,“ Transparency Report, ” (2016).
285. USSIF, "Report on sustainable, responsible and impact investing trends in the United States," *Forum for Sustainable and Responsible Investment*, Washington, D.C. (2018).
286. Vassalou M. and Y. Xing, "Default Risk in Equity Returns," *The Journal of Finance* 59(2)(2004): 831-868. DOI:10.2139/ssrn.297319.
287. Waddock, S. A., and S.B. Graves, "The corporate social performance-financial performance link," *Strategic Management Journal* 18(4) (1997): 303-319.
288. Waddock, S., and S. Graves, "The corporate social performance-financial performance link," (1997).
289. Waldman, D., D. Siegel, and M. Javidan, “Components of CEO transformational leadership and corporate social responsibility,” *Journal of Management Studies* 43(8) (2004): 1703-1725.
290. Windolph, S.E., “Assessing Corporate Sustainability Through Ratings: Challenges and Their Causes,” *Journal of Environmental Sustainability*, 1(1) Article 5 (2011).

291. Yang, Ruoke, "Credit Ratings in the Age of Environmental, Social, and Governance (ESG)," (2020): https://ssrn.com/abstract=3595376 or http://dx.doi.org/10.2139/ssrn.3595376.
292. Yermack, (D. "Flights of Fancy, Corporate Jets, CEO Perquisites, and Inferior Shareholder returns," Journal of Financial Economics 80(1)(2006): 211-42.
293. Zerbib, O. D. "The Green Bond Premium," *Working paper, Tilburg University* (2017) https://ssrn.com/abstract=2890316 .
294. Ziegler et al. 2007 Ziegler, A., Schröder, M., Rennings, K.: The effect of environmental and social performance on the stock performance of European corporations. Environ. Resour. Econ. 37, 661–680 (2007)
295. Ziegler, A., M. Schröder, and K. Rennings, "The effect of environmental and social performance on the stock performance of European corporations," *Environmental Resource Economist,* 37, (2007): 661–680.
296. Zwiebel, J. (1995), “Corporate conservatism and relative compensation”, Journal of Political Economy, Vol. 103 No. 1, pp. 1-25.