Edited by
Marielle de Jong · Dan diBartolomeo

Risks Related to Environmental, Social and Governmental Issues (ESG)

palgrave macmillan

Risks Related to Environmental, Social and Governmental Issues (ESG)

Marielle de Jong • Dan diBartolomeo
Editors

Risks Related to Environmental, Social and Governmental Issues (ESG)

Previously published in *Journal of Asset Management*, Special Issue “Risks related to environmental, social and governmental issues (ESG)”, Volume 22, Issue 2, March 2021

palgrave
macmillan

Editors
Marielle de Jong
Paris, France

Dan diBartolomeo
Boston, MA, USA

ISBN 978-3-031-18229-7

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Journal of Asset Management (2021) 22:77–78
https://doi.org/10.1057/s41260-021-00214-4

EDITORIAL

Introductory editorial

Marielle De Jong[1] · Cécile Diana[1] · Julie Malbois[1]

Published online: 18 March 2021

As soon as he took office, on the 20th of January 2021, Joseph Robinette Biden made the choice to re-join the Paris Climate Agreement. This choice sent the issue of climate change straight back up onto the top of the world agenda. Initiated in 2016 at the COP 21 summit in Paris and since ratified by more than 190 countries, the commitment was made to curb the rise in temperature between 1.5 and 2 degrees Celsius. Or else, survival may be at stake. The commitment calls, as challenging as indeed it sounds, for the building of an economy that the planet can sustain.

It is important to underline that the investment community has a crucial role to play in this endeavour. Since it is up to investors in the end to decide which enterprise to finance and which one not—within the limits of market laws that is—they are *de facto* at the frontline of building the new economy. Market mechanisms have over time, in coherence with Friedman's (1962) doctrine of liberal capitalism, led to investment selection processes that favour economically viable enterprise. The key question today is whether the same market mechanisms can lead to a selection of enterprise that is sustainable as well. Would that be conceivable? Would investors eventually not prefer sustainable over unsustainable, all else equal?

Finance theory tells us that all depends on whether investors judge the gains to weigh up against the risks. That being so, making such judgements is by no means trivial. As a starting point, investors will need access to adequate information, extra-financial information, to be in the capacity to assess the sustainability costs and risks. Yet beyond the new information, it requires new expertise as well. Investors need to work out how to blend the extra-financial data into the regular information flows onto which the existing investment processes rely. The blending must be such as to be able to assess the economic viability and the sustainability of business enterprise in the same time.

✉ Marielle De Jong
marielle.de-jong@grenoble-em.com

1 Grenoble Ecole de Management, Grenoble, France

Adding the extra risk dimension calls for a complete overhaul of the portfolio management process. While new regulation is burgeoning, meaning to streamline the extra-financial information and make it easier to use, the proliferation of this information, extracted from many sources all bundled together into ESG scores, has made the integration of these factors more and more complex. The portfolio manager encounters great difficulties, right from the construction of the portfolio, in quantifying and comparing the extra-financial data. For the information to translate into investment risk, it must be brought back to the investment horizon. Sea levels rise over decades … an Eco tax meant to fight this may come tomorrow. Indeed, a short-term investor will not have the same tack on climate change, for example, as a long-term investor. Integrating all these considerations into the portfolio management methodologies, is what makes sustainable investing challenging, and fascinating.

It is, despite all, arguably a positive point that the risks related to climate change, and more generally to the societal challenges we are facing today, are becoming more perceptible. The more these risks are seen to latently inflict material damage to invested capital, the more they will become part of the investment selection processes. The earlier such transformation takes place, the earlier may asset prices react. Logically, divestments from assets that are seen to finance unsustainable enterprise will push their prices down. This pricing pressure sends out clear signals to firms to clean up their production lines, and with a bit of luck this may happen before the latent risks become real.

It is a time race. The transformation of the investment industry towards one that finances a sustainable economy seems underway. The question is what will go faster: global warming or the corrective action driven, in large part, by the capital markets. Crucial in this race is that investors gain experience in what-is-called sustainable investing. It is work in progress. It is encouraging to see that serious efforts are going into ESG research, and we are thrilled to contribute to these efforts via this special issue.

In the lead article, Frank Fabozzi, Peck Wah Ng and Diana Tunaru study the impact of Corporate Social

Responsibility (CSR) on Corporate Financial Performance (CFP) and credit ratings in Japan. Their findings are mixed. On an aggregated data level, the impact of ESG scores, used as proxies for CSR, on CFP appears negative, whereas on a more granular data level, test results diverge depending on how CSR and CFP are measured. The impact is negative for the accounting values of firms yet turns positive for market values. Moreover, much depends on the individual E, S and G pillars of the ESG scores that are used and on the test method counting in nonlinear effects or not. The authors believe that two opposing effects are at play, stemming from the agency problem, which pushes toward lower CFP, and from the value-enhancing view leaning toward higher CFP. The impact on credit ratings is more convincingly positive. Overall, it is fair to say that Japanese firms are relatively new to corporate social responsibility; according to the Milken Institute, the portion of investors' reports using ESG scores is 18% in Japan in 2018, compared to 39% in the USA and 46% in the EU.

In their article named "Green Bonds: Shades of Green and Brown" Moritz Immel, Britta Hachenberg, Florian Kiesel and Dirk Schiereck make an account of how the green bond market fares today, thirteen years into its existence. Much has happened since the World Bank issued the first Green Bond in 2008. The authors give evidence that green bonds are trading at a (small) premium compared to non-green (brown) bonds. Interestingly, among the green bonds those scoring high on ESG criteria, in terms of Environmental-, Social- and Governance issues, are more expensive than those that do not.

In his article named "Air Pollution, Investor Sentiment, and Excessive Returns" Matthew Muntifering gives evidence of a remarkable market phenomenon. In the same way that stock markets tend to be upbeat on sunny days, he finds an opposed effect coming from air pollution: polluted air in New York City makes the stock markets downbeat. Matthew, who is doing his PhD in the Department of Agricultural Economics and Rural Sociology, Auburn, conducted his study using the air quality index provided by the Environmental Protection Agency.

In their article named "Sustainability Efforts, Index Recognition, and Stock Performance" Moonsoo Kang, K.G. Viswanathan, Nancy A. White and Edward J. Zychowicz study the price behaviour of stocks entering the North America Dow Jones Sustainability Index, a flagship stock index based on ESG criteria that was launched in 1999. The authors find that stock prices go up on the news of entrance, indicating that the selected stocks are in effect in demand. The authors have verified that the positive reaction is not a result of selection bias other than the ESG criteria.

In his article named "Expected and Realized Returns on Stocks with High and Low ESG Exposure" Olaf Stotz shows that there is a discrepancy between the expected return on stocks with high ESG scores and the eventual outcome. Interestingly, the *ex post* realised returns on high-scoring ESG investments largely outdo the *ex ante* expectations based on the financial fundamentals of the underlying firms, that is, over the period from 2008 to 2018 in the US. Olaf attributes this finding to the news on discount rates, which in effect indicates that high ESG assets are in demand.

Journal of Asset Management (2021) 22:79–95
https://doi.org/10.1057/s41260-021-00204-6

ORIGINAL ARTICLE

The impact of corporate social responsibility on corporate financial performance and credit ratings in Japan

Frank J. Fabozzi[1] · **Peck Wah Ng**[1] · **Diana E. Tunaru**[2]

Revised: 11 January 2021 / Accepted: 11 January 2021 / Published online: 12 April 2021

Abstract
We investigate the impact of companies' sustainability efforts on their corporate financial performance (CFP) and credit ratings in Japan, based on a new proxy for corporate social responsibility (CSR)—Sustainalytics' quantitative Environment, Social and Governance (ESG) ratings. We find weak evidence of a negative impact of ESG scores (on an aggregated basis and disaggregated basis) on several accounting measures of CFP. Our quantile regression results reveal a nonlinear pattern across the quantiles, with CSR effects intensifying at the extremal quantiles. However, we find a weak positive relationship between ESG and stock market-based measures, as well as between ESG and credit ratings. Our findings suggest that investors, credit rating agencies (CRAs) and regulators should differentiate between the three types of ESG screening as they interact and contribute in their specific way to the aggregate ESG effect.

Keywords Corporate social responsibility · Corporate financial performance · Credit ratings · Environment, social and governance ratings · Quantile regression

JEL Classification G39 · Q50 · C21 · C23

Introduction

The subject of Corporate Social Responsibility (CSR) or Environment, Social and Governance (ESG) (both terms are used interchangeably in this paper) has gained increasing prominence in the financial community throughout the world as responsible business models are at the core of the transition to a sustainable global economy. This trend is also present in the Asia-Pacific region, as companies are becoming significantly more ESG responsive (Auer and Schuhmacher 2016).

One of the first studies to offer support for CSR primarily based on stakeholder theory is Freeman (1984), who asserted that a firm's management should formulate corporate policies to satisfy not just shareholders, but also other stakeholders such as customers, employees, suppliers, community groups and governments. In addition to traditional financial measures, stakeholders require that managers also disclose performance in terms of CSR. Numerous corporations around the world have already embedded sustainability principles into their business models, while the world's major exchanges have developed sustainability indexes and set minimum standards for sustainability disclosure as a prerequisite for listing companies on their exchanges. While corporate reputation is the main driver in pursuing sustainability efforts, more and more companies worldwide report their CSR activities, as they are increasingly aware of their additional operational and growth benefits. KPMG (2011) found that in 1996 only 300 firms worldwide produced CSR reports, while by 2014 their number increased to more than 7000 worldwide (Khan et al. 2016).

The interaction between (CSR) and corporate financial performance (CFP) has been extensively examined in numerous theoretical and empirical studies. The findings are still to reach consensus as two contrarian approaches have been put forward. On the one side, Milton Friedman (1970, p. 126) contends that in a free society "there is one and only one social responsibility of business—to use its resources and engage in activities designed to increase its profits so

✉ Frank J. Fabozzi
frank.fabozzi@edhec.edu

1 EDHEC Business School, Roubaix Cedex 1, France

2 University of Kent, Kent Business School, Canterbury, Kent CT2 7NZ, UK

long as it stays within the rules of the game, which is to say, engages in open and free competition without deception or fraud." Friedman would view expenditures for CSR as being an illegitimate waste of resources that is in conflict with a firm's responsibility to its shareholders. According to this view, CSR initiatives by corporate management would result in a lower CFP and a lower credit rating and ultimately higher borrowing costs. On the other opposite side, advocates of policies by management directed at CSR (e.g., Barnett and Solomon 2012; Epstein and Rejc-Buhovac 2014) argue that shareholders and creditors will reward the firm with lower funding cost and higher CFP over time.

Given the current global economic agenda, numerous initiatives recommend that institutional investors consider CSR policies in making allocation decisions. For example, in the European Union, the regulatory authorities are considering making it mandatory for institutional investors to include ESG as part of their fiduciary duty. Although in the Asia-Pacific region CSR investing is largely at a nascent stage, CSR is gaining momentum as sovereign and pension funds are increasingly committing to socially responsible investments. Consequently, with the mandatory requirements for institutional investors to include ESG as part of their fiduciary duty, corporate management in the Asia-Pacific region cannot overlook CSR any longer. Investors who do not consider ESG risks in their portfolios may also risk breaching their fiduciary duty (Ottawa 2018).

Credit rating agencies (CRAs) are concerned with ESG issues, which can negatively affect a firm's financial position and leave creditors vulnerable to significant losses (Fitch Ratings 2004). The "Statement on ESG in Credit Risk and Ratings" (Principles for Responsible Investment 2016) calls for CRAs and investors to recognize the importance of considering ESG factors in credit risk analysis and the imperative of making this information transparent.

In this paper, we explore (at both aggregated and disaggregated levels) three ESG aspects for the Japanese market. First, we investigate the impact of CSR (using the Sustainalytics' ESG Rating database) on the Tobin's Q measure. Second, we investigate the impact of CSR on the accounting-based measures of CFP, namely, return on assets (ROA) and return on equity (ROE). Third, we investigate the effects of CSR on credit ratings of Japanese corporations and examine the disaggregated impact of each of the different subscores of ESG on corporate credit ratings.

Following previous indications of a curvilinear CSR–CFP relationship (see Barnett and Solomon 2012), we extend our analysis beyond the ordinary least squares (OLS) regression and try to measure the impact of ESG on different segments of the distribution of the CFP by employing the quantile regression estimation method. Our empirical results provided by the two regression techniques are different with respect to some covariates, suggesting that investors, CRAs and regulators should differentiate between the three types of ESG screening as they interact and contribute in their specific way to the overall ESG effect. The quantile regression analysis provides evidence of a nonlinear CSR–CFP relationship which can be explained by other empirical findings (e.g. Ding et al. 2016) indicating that the relative position of the firm within its specific industry may play an important role in the dynamics of this relationship.

Theory and empirical evidence on the impact of CSR

There are several theories about CSR and its impact on firm valuation based on various metrics of financial performance. On the empirical side, an overwhelming number of studies on the impact of CSR provide mixed evidence leaving the debate unresolved.

Value-Enhancing and agency perspectives: CSR and CFP

There are two general views in the CSR literature, namely the value-enhancing view and the agency view. The CSR value-enhancing view, or the risk mitigation view, asserts that socially responsible firms which help protect the environment, promote social equality and improve community relationships can adhere to value-maximizing corporate governance practices (Ferrell et al. 2016).[1] Several studies link CSR expenditures to future CFP through specific channels such as attracting and retaining high-quality employees, improving the effectiveness of the marketing of products and services, increasing the demand for products and services and providing superior access to valuable resources. Proponents of CSR also identify indirect channels through which CSR expenditures may improve a firm's CFP, including providing a form of reputation insurance and mitigating the likelihood of negative regulatory or legislative action. Still other studies have focused on the individual components of CSR and how they influence borrowing costs and performance.

Benefits of CSR could extend beyond traditional measures of CFP. The recent relevant literature supports a positive stance for CSR. Nguyen et al. (2017) argue that CSR activities can create shareholder value as long as managers are properly monitored by long-term investors who can ensure that managers choose the amount of CSR that maximizes shareholder value. Fatemi et al. (2015) find that

[1] Lys et al. (2015) refer to this as the "investment hypothesis" as current CSR expenditures lead to improvements in future firm performance.

CSR expenditures create value for the firm. Other studies find that voluntary environmental quality is associated with firm value through both the cash flow and the cost of equity components and that ESG strengths increase firm value, while ESG concerns decrease it. In analyzing the impact of the different components of the firm's ESG score, environmental strengths increase the firm's valuation; however, neither social nor governance strengths increase the firm's valuation. Weaknesses in the different components affect (reduce) the firm's valuation in the same way. Klapper and Love (2004) find that better corporate governance is highly correlated with superior operating performance and market valuation for the firms in emerging markets and that firm-level governance is lower in countries with weaker legal systems.

In contrast, the agency view as advocated by Ferrell et al. (2016) generally considers CSR as a managerial agency problem and a waste of corporate resources. Several studies found a mixed or negative relationship between CSR and CFP. According to Bénabou and Tirole (2010) and Cheng et al. (2016), critics of CSR contend that CSR expenditures are an inefficient use of corporate resources and argue that CSR is often a manifestation of managerial agency problems inside the firm. Krüger (2015) argues that socially responsible firms tend to suffer from agency problems as managers engage in CSR that benefits themselves at the expense of shareholders. Moreover, managers engaging in time-consuming CSR activities may lose focus on their core managerial responsibilities (Jensen 2001).

The empirical evidence on the benefits of CSR for US corporations is inconclusive, although predominantly supporting a positive stance on CSR (Margolis et al. 2009). For non-US firms, Xie et al. (2017) find that CSR has no impact on financial performance of firms in China and Vietnam, but that CSR efforts can help companies improve their financial performance only through improving customer satisfaction. Focusing on firms in sensitive industries from BRICS (Brazil, Russia, India, China and South Africa) countries, Garcia et al. (2017) find that the profitability of a firm's assets is negatively associated with only one of the ESG scores, the environmental performance score.

Offering a different perspective, Lys et al. (2015) document that CSR expenditures are not a form of corporate charity, nor do they improve future financial performance. They argue that firms should undertake CSR expenditures only when they anticipate stronger future financial performance and that corporate accountability reporting is another channel through which outsiders may infer insiders' private information about firms' future financial prospects.

Studies investigating the relationship between CSR and CFP generally measure financial performance using either an accounting-based measure of profitability (Aupperle et al. 1985) or a measure of firm stock market performance (Alexander and Buchholz 1978; Vance 1975). For those studies using accounting-based measures, the meta-analysis of Boaventura et al. (2012) revealed that most studies (48%) use return on equity to measure CFP, followed by return on assets (29%). Tobin's Q was used in only 10% of the studies. Studies that use accounting profitability as a measure of CFP find mixed evidence on the link between CSR and CFP, but overall the empirical literature points toward a positive relationship between CSR and CFP (Erhemjamts et al. 2013; Rodgers et al. 2013).

Risk mitigation and agency perspectives: CSR and credit rating

Other ESG-related research addresses the impact of CSR on a firm's costs of financing and stock returns, providing also mixed evidence. During the 2008–2009 financial crisis, Lins et al. (2017) observed that firms with high social capital measured as CSR intensity had stock returns four to seven percentage points higher compared to firms with low social capital. Focusing on responsible practices related to employees, environment and products, El Ghoul et al. (2011) find that responsible US firms experience a lower cost of capital and thus higher valuation. Menz (2010) reports a weak positive relationship between CSR and bond spreads for European firms. Chava (2014) documents that there is an observed positive relationship between expected stock returns and a firms' environmental concerns, and Goss and Roberts (2011) find that firms with below-average environmental and social performance are associated with a higher premium on their cost of private bank debt. In contrast, Sharfman and Fernando (2008) find that firms with good environmental performance have higher leverage and must pay higher bond yields.

At the theoretical level, there are two opposing perspectives regarding the potential impact of CSR initiatives on credit ratings—the risk mitigation (value enhancing) and the agency perspectives. The risk mitigation perspective suggests that CSR activities improve credit ratings. Arguments in favor of CSR center on the negative correlation between CSR and risk. Godfrey (2005) argues that firms with more CSR engagement are exposed to a lower degree of risk. If the investments in CSR lead to lower risk, credit ratings would improve because they provide information about a firm's default probability. Credit rating agencies and debtholders concentrate considerably more on downside risk when reviewing a firm because their payoff on the upside is limited. Consequently, the risk mitigation view suggests that more socially responsible firms are assigned more favorable credit ratings. Empirically, Jiraporn et al. (2014) found that increasing the CSR by one standard deviation results in an improvement of up to 4.5 % in the firm's credit ratings.

On the other hand, the agency view (Jensen and Mecking 1976) argues that CSR investments represent a misallocation of resources, with managers overinvesting in CSR for private benefits instead of maximization of shareholder wealth. It also suggests that by recognizing the agency conflict engendered by CSR efforts, credit rating agencies will assign lower credit ratings to firms with higher CSR. However, the empirical results are mixed. In a recent study, Lioui and Sisto (2017) show that firms highly rated along CSR dimension see their cost of capital increased by 268 basis points.

Datasets

Sample selection

To investigate the relationship between CSR and CFP and between the CSR and credit ratings in Japan, we use data from the following sources: (1) Sustainalytics' ESG Rating database which provides companies' ESG scores based on a range of core and sector-specific indicators; (2) credit ratings from Japan Credit Rating Agency (JCRA) database which provides long-term issuer credit ratings; and (3) Bloomberg database which provides financial statement data.

The Sustainalytics' ESG Rating database covers 530 Japanese companies and provides "overall ESG scores" and component scores of the three pillars, namely E, S and G scores. We filtered this universe to remove banks and financial institutions, as the measures of corporate financial performance (ROA, ROE) and the control variables (for example, leverage, price-to-book ratios, and so on) are not directly comparable between banks and corporations.

For the purpose of this study, two samples are constructed. For the first sample (to study the impact of CSR on CFP), we filter for availability of financial information and Sustainalytics' ESG Ratings, resulting in a reduced sample of 430 firms. For the second sample (to study the impact of CSR on credit ratings), we collect data including credit ratings for 182 firms.

Corporate social responsibility

Constructing a truly comparable and representative measure of CSR has been challenging due to the multidimensionality of the CSR and the limited perspective of the firm's CSR through the measurement of a single dimension (e.g., philanthropy) of CSR (Lydenberg et al. 1986; Wolfe and Aupperle 1991). Waldock and Graves (1997, p. 304) highlighted the "need for a multidimensional measure applied across a wide range of industries and larger samples of companies".

In recent years, most research on CSR relies on the dataset provided by MSCI ESG KLD STATS database; others rely on subjective CSR measures such as a questionnaire, forced-choice survey instruments, reputation index or content analysis. Critiques of MSCI ESG KLD STATS data point out that positive and negative social actions should not be combined as they are both empirically and conceptually distinct components (Mattingly and Berman 2006; Chatterji et al. 2009).

This study aims to provide new insights regarding the effects of CSP on CFP and credit ratings by using the Sustainalytics' ESG Rating for Asia corporates for the measurement of CSR, as it provides a comparable score for each company. The Sustainalytics' ESG Rating dataset has not been widely used in the literature, given the fact that the scores for Asia corporates are available only since 2009. To the best of our knowledge, this study will be one of the first to use the Sustainalytics' ESG Rating dataset to study the impact of CSR.

Sustainalytics is a leading provider of ESG and corporate governance research, ratings and analysis to investors covering 11,000 global companies (1759 Asia companies) across 42 sectors. Overall, Sustainalytics' ESG Rating assesses 150 core and sector-specific indicators with an average of 80 indicators for each company. There are an additional ten indicators for controversial events. Compared to the MSCI ESG KLD STATS[2] database which only expanded its coverage from 2013 to include non-US companies, Sustainalytics' database covers Asia corporates from 2009. An added advantage of Sustainalytics' ESG Rating database over MSCI ESG KLD STATS data is that it allows comparison across multiple peer groups using numerical scores.

The Sustainalytics' ESG Rating dataset not only provides the overall ESG score but also the component scores of the three pillars, namely Environment (E), Social (S) and Governance (G). The Sustainalytics' ESG Rating is a quantitative score on a scale of 1–100 based on a balanced scorecard system. The overall ESG score is computed as a weighted average of the three pillars, with variable weights depending on the peer group. The score of each pillar is, in turn, the weighted sum of the scores on the issues belonging to the respective pillar (see "Appendix A").

For the CSR assessment, fiscal year data are drawn from the Sustainalytics' ESG Rating database for companies from Japan covering the period from the third quarter of 2009 (September 30, 2009) to the second quarter of 2016 (March 31, 2016).

Corporate financial performance

In this study, we employ two accounting metrics, ROA and ROE, as measures of CFP. Extracted both from Bloomberg, ROA and ROE are calculated as the trailing 12 months net

[2] The MSCI ESG STATS database was previously known as the KLD STATS database; the latter covered only US publicly traded companies. MSCI ESG STATS expanded its coverage of non-US companies in 2013.

income divided by the average of the beginning and ending balance of total assets (total common equity) for each financial year, respectively.

We have also considered Tobin's Q (a forward-looking measure of market value) as a proxy for CFP. In contrast to the backward-looking accounting measures, the firm's market value depends on growth prospects, sustainability of profits, or the expected performance in the future (Rust et al. 2004). Market measures are less susceptible to different accounting procedures and represent the investor's evaluation of the ability of a firm to generate future economic earnings (McGuire et al. 1988). Tobin's Q is extracted from Bloomberg, which defines Tobin's Q as the ratio of the market value of a firm to the replacement cost of the firm's assets and calculates this ratio as the sum of market capitalization, total liabilities, preferred equity and minority divided by total assets.

Control variables

There are two different sets of control variables for each sample. These data are extracted from Bloomberg on a fiscal year end basis.

Control variables for sample 1 (to study CSR and CFP)

Size, leverage, cash, price-to-book (PTB) ratio and industry have been suggested in previous research (Ullmann 1985; McWilliams and Siegel 2000; Lys et al. 2015) to be factors that affect a firm's performance and CSR. To isolate the effects of the ESG Total score and component scores on CFP, the following control variables are used: sales[3], cash, leverage, PTB ratio, beta, industry and year.

All the variables (except Industry and Year) have been standardized. Firm size is used as a control variable because larger firms tend to adopt the CSR principles more often (Tsoutsoura 2004). Larger firms also gather more attention and receive more pressure to respond to shareholders' demands (Burke et al. 1986). Sales (as proxy for size) is a relevant variable because there is some evidence that smaller firms may not exhibit as much socially responsible behavior as do larger firms (Waddock and Graves 1997). Larger firms may have greater resources for CSR expenditures and, therefore, may attract greater public pressure to engage in CSR-related activities (Lys et al. 2015; Wu 2006; Teoh et al. 1999). Leverage, measured by long-term debt to total assets, is used as a proxy for risk (Waddock and Graves 1997; Tsoutsoura 2004). The level of management's risk tolerance influences its attitude toward activities that have the potential to elicit savings, incur future/present costs or build/destroy markets. Cash, as a proxy for availability of resources to undertake CSR expenditures, is used as another control variable. Cash is an indicator of firm performance, which some suggest enables or gives rise to the external demand for CSR expenditures (Preston and O'Bannon 1997; Campbell 2007). Price-to-book (PTB) ratio which measures the market value over the book value of a listed company is another control variable. Leverage is also included, as stable firms with lower risk generally appear more likely to make CSR expenditures (Cochran and Wood 1984; Orlitzky and Benjamin 2001).

We control for industry and year fixed effects. Industry is included because the variation in environmental impact, growth prospects, disclosure requirements, and regulatory oversight in different industries is expected to affect the level of CSR expenditures (Karpoff et al. 2005; Griffin and Mahon 1997; Spencer and Taylor 1987). There are 37 industry sectors in the sample, and the segmentation of the industries follows that used in the Sustainalytics' ESG Rating database. Industry is determined in the model by 36 dummy variables. Year is determined in the model by dummy variables from zero to seven to denote each of the fiscal years from 1999 to 2006.

Control variables for Sample 2 (to study CSR and credit ratings)

To isolate the effects of the overall ESG ratings, we consider a set of control variables routinely considered in the relevant credit ratings literature:[4] size, coverage ratio, operating profit margin, leverage ratio, capital intensity ratio and beta.

Firm size is used as a control variable because larger firms tend to garner more attention and receive more pressure to respond to shareholders' demands (Burke et al. 1986). Firm size is shown to be positively related to credit ratings in studies (Bhojraj and Sengupta 2003). Larger firms tend to face comparatively lower business and financial risks and are therefore expected to have lower credit spreads and higher ratings (Oikonomou et al., 2014). The same applies to the coverage ratio and margin variables, as firms that are

[3] Both Sales and Cash (as proxies for size) use the logarithm of total sales, and cash and marketable securities, respectively, and have not been scaled to total assets. This is to isolate the effect of the specific control variable as total assets can be viewed as a measure of size too.

[4] The literature concerning credit ratings has documented many firm characteristics that influence credit ratings. Default risk is found to be inversely related to credit ratings (Lamy and Thompson 1998. Other studies (Blume et al. 1998; Bhojraj and Sengupta 2003; Mansi et al. 2004; Ashbaugh-Skaife et al. 2006) control for a set of variables routinely used in studies of credit ratings to isolate the effects of the CSR variable.

more profitable can afford to be more socially responsible according to the agency perspective.

A higher leverage ratio is associated with higher default risk as firms that accumulate more debt may have more difficulties in servicing that debt. Capital intensity is included to control for differences in companies' asset structures, as companies with greater capital intensity present lower risk to debt providers and thus are expected to have higher credit ratings.

Company credit rating

Given that this study examines only Japanese companies, we used credit ratings from a domestic credit rating agency instead of credit ratings from global credit rating agencies. According to Asian Bankers Association (2000), domestic credit rating agencies have a better understanding and insights of local companies and better access to local information. Credit ratings are extracted from the ratings database of JCRA—the only Japanese rating agency that is officially registered in the USA and certified in the EU, assigning credit ratings to more than 200 foreign issuers, in addition to the domestic issuers in Japan.

Following other studies where commercial credit ratings are used,[5] a measure of a company's credit rating is specified by translating its long-term issuer credit ratings compiled by JCRA to an ordinal scale (from 8 to 1) as follows: AAA and AA+ (8), AA and AA- (7), A, A-, and BBB+ (6) BBB, BBB-, and BB+ (5), BB, BB-, and B+ (4), B and B- (3), CCC (2), and CC and C (1).

Sample Construction

The initial sample is constructed from 530 Japanese corporates covered by Sustainalytics. After accounting for all of the missing information, sample 1 is reduced to an unbalanced panel of 1908 yearly observations from 427 firms across 37 sectors for the period covering fiscal year end 2009 to 2016 (up to fiscal year-end March 2016). "Appendix B" shows the industry breakdown by sample.

The sample is well diversified in terms of industry representation, with a total of 37 industries, where the first three industries (Chemicals, Machinery and Technology Hardware) each represent 7% of the sample. For the second sample (to investigate the relationship between CSR and credit ratings), we further filtered the sample to require firms to have credit ratings. Based on these criteria, the sample is reduced to 182 firms. For each firm, fiscal year-end financial data for the period from 2009 to 2015 are collected. Corresponding credit ratings and ESG scores (with a three-month lag from fiscal year end) are extracted from the JCRA database and Sustainalytics' ESG Rating database, respectively.

After accounting for all of the missing information, sample 2 is reduced to an unbalanced panel of 855 observations from a total of 182 firms across 33 industry sectors for the period covering fiscal year end 2009 to 2015. Based on the industry breakdown in "Appendix B", the firms are well spread over the 33 industries, with only one industry (Transportation) accounting for 12% of the sample. Each of the other industries accounts for less than 10% of the sample. The top five industries (Transportation, Chemicals, Food Products, Utilities and Machinery) represent approximately 42% of the sample. Table 1 provides the descriptive statistics of the dependent and independent

Table 1 Descriptive statistics for regression variables—sample 1

Variable	Mean	SD	Min	Max
Tobin's Q ratio	1.3129	0.8553	0.5989	14.0065
ROA (%)	3.7995	4.6347	− 65.2341	36.2296
ROE (%)	7.1138	13.0174	− 197.3558	131.5627
Sales ($)	3.8298	0.5187	2.1108	5.4266
Leverage	21.8723	17.5416	0	72.7965
Cash ($)	2.9889	0.5015	0.5345	4.6383
Price-to-book ratio	1.5805	1.2921	0.4234	18.2719
Beta	0.7705	2.824	− 23.193	14.627
Overall ESG score	0.5605	0.0787	0.32	0.8
Environment score	0.6009	0.1252	0.26	0.93
Social score	0.5287	0.0909	0.27	0.87
Governance score	0.554	0.0664	0.36	0.84

Note: This table provides the descriptive statistics of the examined variables for 427 Japanese firms during the period 2009 to 2016, based on 1908 observations. Overall ESG score represents Sustainalytics' ESG Rating of a company's overall ESG performance on a scale of 1–100 expressed in percentage.

Table 2 Descriptive statistics for regression variables—Sample 2

Variable	Mean	Median	SD
Rating	6.70	7.00	0.72
Overall ESG score	0.5648	0.56	0.0706
Size	4.04	4.02	0.44
Coverage ratio	77.82	9.95	553
Margin	6.94	5.82	6.44
Leverage ratio	22.28	19.14	15.05
Capital intensity ratio	41.44	36.97	20.63

This table provides the descriptive statistics of the examined variables representing 182 firms during the period 2009 to 2015, based on 855 observations. Rating is the long-term issuer credit ratings compiled by JCRA transformed to an ordinal scale that ranges from 1 to 8. Overall ESG score represents Sustainalytics' ESG Rating of a company's overall ESG performance on a scale of 1–100 expressed as percentage.

[5] See Attig et al. (2013), Blume et al. (1998), Bhojraj and Sengupta (2003), Mansi et al. (2004), and Ashbaugh-Skaife et al. (2006).

variables for Sample 1. The mean Tobin's Q, ROA and ROE are 1.31, 3.80% and 7.11%, respectively. With respect to ESG scores, the sample has a mean overall ESG score of 0.56, while the mean E score, S score and G score are 0.60, 0.53 and 0.55, respectively. These scores are reflective of the CSR awareness and integration within firms in Japan. Table 2 provides summary descriptive statistics for Sample 2. The mean credit rating of the firms in the sample (out of a scale of 1 to 8) is 6.7 and the mean overall ESG score is 0.56.

Methodology

Here are the main hypotheses that will be tested in our paper. Without loss of generality, we denote by SCORE one of the following Sustainalytics' ESG scores: overall ESG, the disaggregated E, the disaggregated S or the disaggregated G.

We test the significance of the relationship between the CSR and CFP based on the following hypothesis:

Hypothesis 1 Firms that implement CSR initiatives as measured by SCORE experience a significant change in their financial performance.

For the relationship between the CSR and credit ratings, we formulate the null hypothesis to test for the positive direction suggested by previous empirical studies:

Hypothesis 2 Firms that implement CSR initiatives as measured by SCORE experience a significant and positive change in their credit rating.

Estimation Models for CFP

First, the relationship between CSR and CFP using both accounting measures (ROA and ROE) and the stock market-based (Tobin's Q) measure is tested using a two-way fixed-effects pooled regression model after controlling for the four key financial variables (in lagged terms as proxies for size, leverage, cash holdings and price-to-book ratio) and Beta. The model specification takes into account both fixed industry and time effects by including 36 and six industry and time dummy variables, respectively. Considering that overall ESG scores may hide confounding effects of the different dimensions of CSR, this study also looks into both the overall ESG scores and the disaggregated ESG scores, namely the E score, S score and G score. For each CFP proxy (Tobin's Q, ROA and ROE) as the dependent variable, we estimate the following model in Eq. (1) using pooled OLS and quantile regression estimation methods:

$$CFP_{it} = \alpha + \beta_1 \mathrm{ESG}_{i,t-1} + \beta_2 \mathrm{Sales}_{i,t-1} + \beta_3 \mathrm{Leverage}_{i,t-1} + \beta_4 \mathrm{Cash}_{i,t-1} + \beta_5 \mathrm{PTB}_{i,t-1} + \beta_6 \mathrm{Beta}_{i,t-1} + \sum_{k=1}^{36} a_k \mathrm{ID}_{ik} + \sum_{j=1}^{6} b_j \mathrm{TD}_{ij} + \varepsilon_{it} \tag{1}$$

where Tobin's Q is the ratio of the market value of a firm to the replacement cost of the firm's assets (extracted from Bloomberg); ROA is Return on Assets (extracted from Bloomberg) computed as the trailing 12 months net income divided by the average of the beginning and ending balance of total assets for each financial year; ROE is Return on Equity (extracted from Bloomberg) computed as the trailing 12 months net income divided by the average of the beginning and ending balance of total common equity for each financial year; ESG is a measure of a firm's sustainability performance based on respective overall ESG, E, S and G scores (extracted from Sustainalytics' ESG Rating database); Sales is the logarithm of total sales in US dollars (US$ millions converted at the prevailing exchange rate at the end of each fiscal year); Cash is the logarithm of cash and marketable securities (millions converted at the prevailing exchange rate at the end of each fiscal year); Leverage is the leverage ratio as measured by the ratio of long-term debt to total assets; price-to-book ratio (PTB ratio) is the ratio of a stock's market value over its book value as at each fiscal year end; Beta is the measure of the firm's systematic risk (extracted from Bloomberg) and is computed based on the regression of the historical trading prices of the stock using weekly data over a two-year period; ID is the respective industry dummy variable which reflects the industry segments provided by Sustainalytics' ESG Rating database, and ID denotes the year dummy variable to reflect the respective fiscal year of the financial data.

We also investigate whether the CFP differs across quantiles of the conditional distribution by employing a quantile regression analysis. We briefly explain the main idea behind the quantile regression model introduced by Koenker and Bassett (1978) as an extension of the conditional mean estimation to prediction of conditional quantile for the dependent variable as functions of the independent variables.

If we denote the dependent variable by Y with its distribution function F_Y and the quantile position by $\tau \in (0, 1)$, then the quantile function for the τ^{th} quantile is defined as $q_Y(\tau) = F_Y^{-1}(\tau) = \inf\{Y : F_Y(Y) \geq \tau\}$. This can be interpreted as following: $100\tau^{th}\%$ of the probability mass of Y is below $q_Y(\tau)$.

Each quantile of the conditional distribution of the response variable is expressed as a function of the observed explanatory variables. Considering the following quantile family $\tau = \{0.1, 0.3, 0.5, 0.7, 0.9\}$, the quantile analysis comprises five regression equations

$$Y_{i\tau} = \alpha_\tau + \sum_j \beta_{j\tau} X_{ij} + \varepsilon_{i\tau}$$

where α_τ and $\beta_{j\tau}$ are estimated by minimizing a special objective function equal to the sum of *asymmetrically* weighted absolute residuals (see Koenker and Hallock 2001*)* and not by the OLS method. The group of explanatory variables is the same as in Eq. (1).

The quantile regression allows us to measure potentially changing impact levels of the same explanatory variables as in Eq. (1), on different segments of the distribution of the dependent variable. While the OLS regression analysis provides a best-fit methodology for the mean of the dependent variable, the quantile regression provides a best fit for a specific quantile of the distribution around that mean value. By employing the quantile regression, we avoid some of the issues present within standard OLS regression, more specifically the influence of outliers and dependence on assumptions regarding the residuals. We keep the same set of dependent variables as in the pooled regression and the same treatment to the variables by standardizing them.

For the study on CSR, we examine in a first stage the impact of overall ESG score on credit ratings after controlling for the five key financial variables that are known to affect credit ratings. A probit regression model is used given the ordinal (discrete) nature of the dependent variable (Credit Rating) in line with prior research. This regression approach is used to test whether information on CSR activities (measured by overall ESG score), distinct from information considered by rating agencies, can have explanatory power on a company's credit ratings. In a second stage, we extend the analysis by including dummy variables to measure industry and year fixed effects:

$$\begin{aligned} &\text{probit}(\text{CR}_{it}) = \alpha + \beta_1 \text{ESG}_{i,t-1} + \beta_2 \text{Size}_{i,t-1} + \beta_3 \text{Coverage}_{i,t-1} + \beta_4 \text{Margin}_{i,t-1} \\ &+\beta_5 \text{Leverage}_{i,t-1} + \beta_6 \text{Capital intensity} + \beta_7 \text{Beta}_{i,t-1} + \\ &+ \left[\sum_{k=1}^{32} a_k \text{ID}_{ik} + \sum_{j=1}^{6} b_j \text{TD}_{ij} \right] + \varepsilon_{it} \end{aligned} \tag{2}$$

where *CR* refers to the credit rating of the company, *Size* is the logarithm of total assets in US dollars (millions converted at the prevailing exchange rate at the end of each quarter), *Coverage ratio* is the ratio of earnings before interest and taxes divided by interest expense (EBIT/Interest), *Margin* is the operating profit margin (the ratio of operating income to sales), *Leverage ratio* is the ratio of long-term debt to total assets, *Capital intensity* is the ratio of net fixed assets to total assets and *ESG, Leverage, Beta, ID* and *YD* variables have been previously defined for Eq. (1).

Empirical results

CSR–CFP empirical results

We measure the impact of CSR on three metrics of CFP (Tobin's Q, ROA and ROE) using OLS and quantile regression models. Given that the optimization algorithms involved in the estimation of the two types of regression are different, the estimation results are not directly comparable. However, the new insights provided by the quantile regression are of great value as they suggest relationships of different intensity and sometimes of a different direction between the examined variables, when compared with the results from the pooled OLS regression approach. We collate the results of both types of models in Tables 3, 4 and 5 for Tobin's Q, ROA and ROE, respectively. Each table contains four panels corresponding to the aggregate ESG score and the three individual pillars E, S and G. All the control variables (Sales, Cash, Leverage, PTB ratio, and Beta) were initially included in the OLS regressions to ascertain whether they are potential predictors. The OLS regression results present the estimates of the final specification after the elimination of the insignificant (5%) covariates, such as Cash and/or Sales.

To address year and industry effects, dummy variables are assigned to the different fiscal years (from 2010 to 2016) and the different industry sector (see per industry breakdown in "Appendix B"). The pooled regressions were initially estimated without considering the year and industry effects. With inclusion of year and industry effects, the R-squared generally increased across the different models. The results between the two estimation methods are in general consistent in the case of the two accounting measures ROE and ROA, and less convergent when CFP is measured by Tobin's Q ratio.

For the Tobin's Q measure as the dependent variable (Table 3), the effect of the overall ESG score estimated by the OLS regression is small and positive (0.038) and significant at the 5% level of significance. These findings (positive relationship between ESG and Tobin's Q) are in line with the majority of the literature. The rationale often used in support of CSR improving firm value rests on increased transparency that mitigates information asymmetry between investors and the firm, leading to positive outcomes such as better access to capital.

Table 3 Estimation results CSR–CFP (Tobin's Q) relationship

Regression	Pooled regression	Quantile regression				
Coefficients	OLS	0.1	0.3	0.5	0.7	0.9
Intercept	− 0.038	− 0.340***	− 0.177***	− 0.081***	0.025	0.219***
ESG	0.038**	− 0.007	0.007	− 0.001	− 0.005	− 0.016
Sales	− 0.079***	0.040***	− 0.012	− 0.024***	− 0.038***	− 0.062***
Leverage	− 0.169***	− 0.001	− 0.027***	− 0.039***	− 0.046***	− 0.042**
Cash	–	− 0.019**	0.018**	0.022**	0.035***	0.057***
PTB	0.813***	0.311***	0.528***	0.695***	0.835***	1.144***
Beta	0.089***	0.022**	0.031***	0.040***	0.052***	0.069***
Coefficients	OLS	0.1	0.3	0.5	0.7	0.9
Intercept	− 0.35	− 0.334***	− 0.175***	− 0.081***	0.027	0.216***
E	0.018	− 0.009	0.01	0	− 0.005	− 0.015
Sales	− 0.070***	0.043***	− 0.012	− 0.026***	− 0.039***	− 0.062***
Leverage	− 0.170***	− 0.003	− 0.028***	− 0.039***	− 0.046***	− 0.047***
Cash	–	− 0.021**	0.017**	0.022***	0.036***	0.056***
PTB	0.813***	0.318***	0.531***	0.695***	0.834***	1.13***
Beta	0.900***	0.022**	0.032***	0.040***	0.052***	0.068***
Coefficients	OLS	0.1	0.3	0.5	0.7	0.9
Intercept	− 0.03	− 0.341***	− 0.178***	− 0.080***	0.024	0.218***
S	0.028**	− 0.002	0.005	− 0.002	− 0.005	− 0.011
Sales	− 0.072***	0.038***	− 0.011	− 0.024**	− 0.039***	− 0.059***
Leverage	− 0.168***	− 0.003	− 0.028***	− 0.039***	− 0.045***	− 0.046**
Cash	–	− 0.020**	0.016***	0.021***	0.035***	0.056***
PTB	0.811***	0.315***	0.523***	0.695***	0.836***	1.143***
Beta	0.089***	0.023**	0.032***	0.039***	0.053***	0.068***
Coefficients	OLS	0.1	0.3	0.5	0.7	0.9
Intercept	− 0.002	− 0.337***	− 0.177***	− 0.082***	0.035	0.247***
G	0.042***	0.002	0.003	0.001	0.009	0.006
Sales	− 0.075***	0.038***	− 0.01	− 0.026***	− 0.044***	− 0.062***
Leverage	− 0.170***	− 0.001	− 0.028***	− 0.039***	− 0.044***	− 0.050**
Cash	–	− 0.020**	0.017***	0.023***	0.038***	0.052***
PTB	0.814***	0.315***	0.529***	0.695***	0.831***	1.145***
Beta	0.090***	0.023***	0.032***	0.040***	0.051***	0.069***

The four panels of this table report the OLS pooled and quantile regression results for Tobin's Q (dependent variable) on overall ESG score and the three individual Pillars (Environment Score, Social Score and Governance Score) for Japan (***$p < 0.01$; ** $p < 0.05$; * $p < 0.10$).

The evidence provided by the quantile regression indicates that the coefficients are very small and statistically insignificant across all quantiles.

When we disentangle the overall ESG score into its three individual components E, S and G, the main driving factors suggested by the OLS estimation are G and S, while none of them is significant across all quantile levels. Other divergent effects are present for the control variable Cash, as it is insignificant according to the OLS estimation, but with a clear trend from negative to positive effect across the quantiles in all regressions. The effects of other three control variables are consistent between the two estimation methods, being significant and positive for PTB and Beta, and negative for Leverage. Moreover, the quantile analysis reveals a positive trend in magnitude, as the impact of these three variables on Tobin's Q ratio intensifies as we move towards a higher quantile. Therefore, for the Tobin's Q case based on the OLS regression results, we accept Hypothesis 1 for ESG, S, and G pillars while we reject this hypothesis for the Environment pillar of the ESG. This confirms previous findings (see Bouslah et al. 2010) on the environmental performance suggesting that financial markets have not yet priced in the benefits of such practices. The quantile analysis rejects all

Table 4 Estimation Results CSR–CFP (ROA) Relationship

Regression	Pooled regression	Quantile regression				
Coefficients	OLS	0.1	0.3	0.5	0.7	0.9
Intercept	0.154*	− 0.637***	− 0.202***	0.036	0.236***	0.607***
ESG	− 0.045**	− 0.046	0.004	0.005	− 0.031*	− 0.108***
Sales	− 0.079***	− 0.051	− 0.095***	− 0.074***	− 0.065***	− 0.04
Leverage	− 0.364***	− 0.196***	− 0.224***	− 0.245***	− 0.249***	− 0.267***
Cash	–	− 0.075*	0.012	0.049***	0.075***	0.065***
PTB	0.444***	0.249***	0.392***	0.523***	0.616***	0.824***
Beta	0.055***	0.053*	0.021	0.029**	0.033**	0.044***
Coefficients	OLS	0.1	0.3	0.5	0.7	0.9
Intercept	0.175**	− 0.664***	− 0.201***	0.033	0.249***	0.649***
E	− 0.057**	− 0.003	0.01	0.008	− 0.029	− 0.063*
Sales	− 0.077***	− 0.062*	− 0.099***	− 0.073***	− 0.064***	− 0.054*
Leverage	− 0.362***	− 0.217***	− 0.223***	− 0.245***	− 0.253***	− 0.274***
Cash	–	− 0.102***	0.014	0.045**	0.076***	0.078***
PTB	0.442***	0.248***	0.392***	0.523***	0.607***	0.870***
Beta	0.055***	0.051*	0.021	0.028**	0.030**	0.047***
Coefficients	OLS	0.1	0.3	0.5	0.7	0.9
Intercept	0.146*	− 0.690***	− 0.198***	0.038	0.243***	0.635***
S	− 0.040*	− 0.067**	0.004	0.006	− 0.017	− 0.074***
Sales	− 0.086***	− 0.041	− 0.094***	− 0.075***	− 0.074***	− 0.071***
Leverage	− 0.366***	− 0.178***	− 0.224***	− 0.246***	− 0.249***	− 0.258***
Cash	–	− 0.088**	0.009	0.047***	0.074***	0.076***
PTB	0.446***	0.248***	0.392***	0.522***	0.614***	0.843***
Beta	0.056***	0.033	0.021	0.029**	0.036***	0.042***
Coefficients	OLS	0.1	0.3	0.5	0.7	0.9
Intercept	0.143*	− 0.674***	− 0.213***	0.022	0.228***	0.649***
G	0.018	− 0.025	− 0.013	− 0.005	− 0.015	− 0.047**
Sales	− 0.103***	− 0.065*	− 0.086***	− 0.067***	− 0.075***	− 0.076***
Leverage	− 0.364***	− 0.215***	− 0.224***	− 0.248***	− 0.247***	− 0.272***
Cash	–	− 0.091**	0.009	0.047***	0.074***	0.076***
PTB	0.444***	0.246***	0.388***	0.522***	0.609***	0.869***
Beta	0.055***	0.052*	0.022*	0.031**	0.030**	0.048***

Note: The four panels of this table report the OLS pooled and quantile regression results for ROA (dependent variable) on overall ESG score and the three individual pillars (Environment Score, Social Score and Governance Score) for the Japan. Countries (***$p < 0.01$; ** $p < 0.05$; * $p < 0.10$).

four hypotheses, indicating that there is no significant impact of the ESG efforts at both aggregate and individual levels.

When ROA is used as a proxy for CFP, the empirical results from the two types of regression seem to reconcile, but only at the extremal quantiles (see Table 4). The general conclusion is that the overall ESG and individual E and S scores have a negative impact on the ROA measure. The difference between the two regression types concerns the Governance pillar, which is insignificant in the OLS analysis and negative and significant in the quantile regression at the 0.9 quantile. This particular pattern is observed across the quantiles also in the coefficient estimates for the overall ESG and the individual E and G pillars. In other words, companies with the highest ROA seem to be at a financial disadvantage if they try to satisfy the ESG criteria. Moreover, for the Social pillar the observed effect is more complex, exhibiting a nonlinear dependence. More specifically, although like in Barnett and Solomon (2012) we find that the ESG–CFP relationship (through the S pillar) has a U shape, in our study it is an inverse shape as the negative effect intensifies at both extremal quantiles. According to the OLS analysis, among the individual pillars, the Environmental pillar has the most negative impact (− 0.057) which is higher than the aggregate ESG effect (− 0.045). For the quantile analysis, the driving individual factor is the Social one (− 0.067 at the 0.1 quantile and − 0.074 at the 0.9 quantile). The Leverage and PTB covariates have a consistent positive trend across the quantiles with both negative and positive effects that intensify as the quantile level increases. The results concerning the Cash control variable show an insignificant coefficient in the OLS regression and a significant changing sign from negative to positive in the quantile regression. The results of the quantile regression are more realistic as they correctly identify that companies with an inferior financial performance do not benefit from increasing their cash position, while well-performing firms do. According to both regressions, we accept Hypothesis 1 at the 10% level of significance for the ESG, E and S scores. Hypothesis 1 of a significant relationship between CSP and CFP through the Governance pillar is rejected for the OLS regression, but accepted for the quantile regression at the 0.9 quantile.

When CFP is measured by ROE, the results are similar to those when employing ROA. The evidence presented in Table 5 shows that the OLS regression results are mixed, as the relationship between ROE and CSR is negative and significant (at the 1% level) at the aggregate ESG level and S pillar level, insignificant for the G score and positive and significant for E score. The coefficients for ESG and S scores are about

Table 5 Estimation Results CSR–CFP (ROE) Relationship

Regression	Pooled regression	Quantile regression				
Coefficients	OLS	0.1	0.3	0.5	0.7	0.9
Intercept	0.178*	− 0.422***	− 0.091***	0.100***	0.278***	0.563***
ESG	− 0.097***	− 0.056*	− 0.021**	− 0.019**	− 0.024**	− 0.060***
Sales	–	0.004	0.043***	0.068***	0.087***	0.134***
Leverage	− 0.192***	− 0.157***	− 0.060***	− 0.028***	− 0.001	0.042*
Cash	–	− 0.042	− 0.035**	− 0.022*	− 0.01	− 0.01
PTB	0.227***	0.184***	0.290***	0.386***	0.463***	0.576***
Beta	0.063***	0.024	0.017*	0.026***	0.028***	0.058***
Coefficients	OLS	0.1	0.3	0.5	0.7	0.9
Intercept	0.214**	− 0.392***	− 0.082***	0.111***	0.284***	0.586***
E	0.107***	− 0.029	− 0.014	− 0.014*	− 0.013	− 0.038***
Sales	− 0.103***	− 0.005	0.040***	0.062***	0.083***	0.119***
Leverage	0.188***	− 0.172***	− 0.061***	− 0.032***	− 0.003	0.049**
Cash	–	− 0.037	− 0.035***	− 0.018	− 0.014	− 0.009
PTB	0.222***	0.194***	0.291***	0.378***	0.461***	0.580***
Beta	0.063***	0.033	0.020**	0.028***	0.027***	0.059***
Coefficients	OLS	0.1	0.3	0.5	0.7	0.9
Intercept	0.159	− 0.444***	− 0.099***	0.099***	0.277***	0.538***
S	− 0.086***	− 0.071**	− 0.014	− 0.012	− 0.018*	− 0.049***
Sales	–	0.026	0.035***	0.067***	0.082***	0.132***
Leverage	− 0.199***	− 0.152***	− 0.055***	− 0.028***	0	0.047**
Cash	–	− 0.052	− 0.031**	− 0.024**	− 0.011	− 0.016
PTB	0.234***	0.188***	0.290***	0.384***	0.463***	0.572***
Beta	0.064***	0.014	0.019*	0.028***	0.030***	0.052***
Coefficients	OLS	0.1	0.3	0.5	0.7	0.9
Intercept	0.137	− 0.398***	− 0.078***	0.103***	0.268***	0.540***
G	0.006	− 0.03	− 0.020**	− 0.022***	− 0.023***	− 0.048***
Sales	–	− 0.007	0.037***	0.070***	0.088***	0.115***
Leverage	− 0.204***	− 0.170***	− 0.061***	− 0.036***	− 0.003	0.042*
Cash	− 0.059**	− 0.046	− 0.031**	− 0.025**	− 0.015	0.004
PTB	0.231***	0.195***	0.295***	0.387***	0.461***	0.590***
Beta	0.061***	0.037	0.021**	0.031***	0.028***	0.059***

Note: The four panels of this table report the OLS pooled and quantile regression results for ROE (dependent variable) on overall ESG score and the three individual pillars (Environment Score, Social Score and Governance Score) for the Japan. Countries (***$p < 0.01$; ** $p < 0.05$; * $p < 0.10$).

– 0.1 and are higher than the regression results for ESG and ROA. However, the R-squared values are lower at about 14%.

Moving to the quantile regression, the results are consistent across all four regression models, with a negative and significant impact of all ESG scores, at both aggregate and individual levels. The ESG overall effect and the Governance effect are uniformly spread across the quantiles, while the Environment and Social pillars have a significant impact only at the extremal quantiles, 0.1 and 0.9, respectively. Again, the concave U-shape pattern is clearly present in the Social component, an effect that still exists but diminishes for the overall ESG score. This suggests that we should differentiate between the three types of ESG screening as they interact and contribute in their specific way to the overall ESG effect. With regard to the control variables, Leverage is predominantly negatively correlated with CFP, while for Sales, Beta and Cash holdings, the correlation results produce mixed evidence of significant results. In the case of price-to-book ratios, the relationship with all measures of CFP is positive and significant. For both estimation methods, we accept Hypothesis 1 for ESG, E, and S, whereas the same hypothesis for the G pillar is rejected in the case of the OLS regression but it is accepted based on the quantile regression.

The two regression analyses above present us with various results evidenced by the CSP-CFP literature: negative, positive or no significant relationship. There are several potential reasons for these findings.

First, CSR expenses have the potential to drain the firm's resources and reduce its immediate cash flows and profitability as evidenced by the negative impact on ROA and ROE. Second, this general lack of a significant positive relationship between ESG and CFP possibly occurs because the companies earmark part of their investments for environmental practices, thereby failing to allot them to the companies' profitable activities. This, in turn, could stem from the relatively higher costs of CSR expenditures to comply with government- and nongovernment-imposed corporate ESG disclosures guidelines.

However, the results from the pooled regressions show that there is a gain in firm value as measured by Tobin's Q from CSR efforts (based on overall ESG scores and disaggregated ESG scores). This positive and significant relationship shows that better alignment of corporate strategies with social responsibility initiatives may generate higher levels of firm value observed in the data. Nguyen et al. (2017) also find that positive valuation interaction between CSR and shareholder value is not driven by higher profitability but by lower cash flow risk—via better stakeholders' relations, lower likelihood of legal actions and greater customer loyalty. Moreover, the long-term benefits of CSR efforts (improving probability of survival, lengthening the longevity of its cash flows or lowering its cost of capital) can outweigh the costs and improve market value.

Table 6 Probit Regression results on the effect of overall total score and individual pillar scores on credit ratings

	Overall ESG	E	S	G
Overall ESG score	0.048***	–	–	–
E score	–	0.032***	–	–
S score	–	–	0.020***	–
G score	–	–	–	0.036***
Size	0.607***	0.597***	0.765***	0.771***
Coverage ratio	0.001**	0.001*	0.001**	0.001*
Margin	0.052***	0.050***	0.045***	0.045***
Leverage ratio	– 0.025***	– 0.024**	– 0.029***	– 0.029***
Capital intensity ratio	0.0361***	0.038***	0.035***	0.034***
Beta	0.003	0.002	0.003**	– 0.002
Pseudo-R-squared	7.76%	7.87%	7.05%	7.27%

This table presents results of ordered probit regressions of companies' credit ratings with the Overall ESG scores and individual pillar scores as the target independent variables, respectively. Based on one-tail test, ***, ** and * denote statistical significance at 0.1%, 1% and 5% levels respectively

The lack of a statistically significant relationship could be partially attributable to mandatory regulations in place for ESG so that the market does not reward CSR efforts. According to the Global Guide to Responsible Investment Regulation (PRI 2016) which laid out the Regulation Map Summary, Japan appears to have relatively stringent disclosure guidelines.

Overall, our results illustrate various aspects that call for more in-depth consideration when one explores how CSR initiatives impact a firm's financial performance. The empirical evidence shows that the relationship between CSP and CFP may depend on the proxy we use for CFP. We may invoke here a temporality issue, as market-based measures (Tobin's Q) are long-term metrics, while the accounting measures (ROA and ROE) are short term. We bring new empirical evidence that different estimation methods can yield contradictory conclusions with significant long-term consequences for all the stakeholders. The consistency of the results produced by the quantile regression makes this technique superior to the OLS estimation and allows us to form a conclusion that supports a negative CSR effect on the financial performance of Japanese firms.

CSR-Credit Ratings Empirical Results

The results of the ordered probit regression are presented in Table 6. There are four regression models corresponding to the second hypothesis, where the target covariates are overall ESG, E, S and G scores, respectively.

Table 7 Probit regression results on the effect of overall ESG score and individual scores on credit ratings including Industry and Year Effects

	Overall ESG	E	S	G
Overall ESG score	0.069***	–	–	–
E score	–	0.061***	–	–
S score	–	–	0.010	–
G score	–	–	–	0.053***
Size	1.021***	0.896***	1.335***	1.152***
Coverage ratio	0.000	0.000	0.000	0.000
Margin	0.065***	0.071***	0.054***	0.054***
Leverage ratio	− 0.068***	− 0.065***	− 0.072***	− 0.076***
Capital intensity ratio	0.061***	0.058***	0.058***	0.060***
Beta	0.004	0.000	0.008	0.002
Industry effects	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
Pseudo-R squared	20.45%	21.34%	19.37%	19.99%

This table presents results of ordered probit regressions of companies' credit ratings with the overall ESG scores and Individual pillar scores as the target independent variables and double fixed effects. Based on one-tail test, ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively

The results indicate that Hypothesis 2 is supported. The effect of the overall ESG score on credit ratings is positive and statistically significant at the 1% level, which is consistent with most previous studies. The evidence suggests that firms with higher overall ESG scores enjoy better credit ratings. The estimation results also show that the individual E, S and G scores are positively correlated with credit ratings, supporting the risk mitigation view (positive association between CSR activities and credit ratings) over the agency view (negative relationship between CSR activities and credit ratings).

For the control variables, the coefficient for the size variable is positive and significant at the 0.1% level, confirming that larger firms seem to have lower risk of default. Likewise, for the Margin variable, the operating margin is positively correlated with ratings because higher profitability is associated with lower default risk. For the coverage ratio, the correlation is weak—higher interest coverage is positively correlated with ratings only at the 10% level. The estimated coefficient on the Leverage ratio is negative as firms that have higher debt have lower credit ratings or higher default probability. A positive coefficient for Capital Intensity is consistent with expectations that companies with greater capital intensity present lower risk to debt providers, and thus they are expected to have higher credit ratings. For Beta, there are no significant results.

These results support the risk mitigation view and suggest that there is a significant relationship between credit ratings and both overall ESG score (which is an aggregation of different pillars of ESG), as well as scores of the disaggregated pillars of ESG for the Japanese companies in our sample.

Industry and Year Effects

To obtain a further understanding of the relationship between credit ratings and ESG scores, we augment the analysis by considering year effects and industry effects. Dummy variables are assigned to the different fiscal years (from 2009 to 2015) and the different industry sectors (as per industry breakdown in "Appendix B").

Table 7 presents results of probit regressions of companies' credit ratings on the overall ESG scores and individual pillar scores for Hypothesis 2 with the addition of these dummy variables. Compared to the results reported in Table 6, the overall ESG scores as well as the E and G scores are positively correlated with credit ratings, but the effects of the S scores on credit ratings are not statistically significant after taking into account industry and year effects. The positive correlation found between the individual E and G scores and credit ratings in this study survives this robustness check and suggests that heightened efforts on environment and governance issues would have a statistically significant impact on credit ratings. This is particularly pertinent considering that approximately 36% of the sample is from environmentally sensitive industries (the top four industries in the sample—Transportation, Chemicals, Food Products and Utilities). These results emphasize again the importance of disaggregating the overall ESG scores which may, on an aggregate basis, hide confounding effects among the different pillars of CSR. The coefficients for Size, Margin and Capital Intensity are all positive and significant except for coverage ratio and Beta with no significant results). Conversely, the estimated coefficient on the Leverage ratio remained negative.

With the inclusion of year and industry effects, Hypothesis 2 is supported for ESG, E and G, while it is rejected for the S pillar. The R-squared increased to above 20% compared to the R-squared of the results without taking into account industry and year effects (of about 7%).

The results illustrate that credit ratings have implicitly considered CSR strengths and weaknesses in addition to financial parameters. While firms with more CSR engagement are generally exposed to a lower degree of risk or better credit ratings, these findings isolate the two pillars (E and G) in ESG that impact credit ratings. These results are in line with those of Ashbaugh-Skaife et al. (2006) who present evidence that firms exhibiting stronger corporate governance (with attributes such as higher degree of financial transparency, board independence, board expertise and the like) benefit from higher overall firm credit ratings. Similarly, Ge and Liu (2015) report that bondholders are more likely to use CSR performance information to assess the creditworthiness of issuers with weaker corporate governance and those operating in environmentally sensitive industries.

Conclusions

The OLS estimation results suggest a positive impact of aggregated CSR on CFP (as measured by Tobin's Q), while there is significant evidence of a negative correlation between CSR and CFP (as measured by ROA and ROE). Although the empirical evidence from the quantile regression analysis is in general similar to the OLS results, the negative association is present across all three proxies considered for CFP, including the Tobin's Q measure. These findings support the agency theory that the managers of nonfinancial Japanese companies consider as their main target the maximization of shareholders' wealth, a pattern also prevailing among nonfinancial Chinese companies (see Farag et al. 2015). Moreover, we have identified a pattern of significance, as the negative ESG impact seems to exist and intensify only across the extremal quantiles, especially at the 0.9 percentile level.

These findings support the potentially nonlinear characteristic of the CSR–CFP relationship suggested by Sahut and Pasquini-Descomps (2015). Our analysis across quantiles shows that for Japanese companies with medium financial performance there is no evidence of a significant ESG impact, while companies in a strong and sometimes weak financial position are negatively affected by increasing efforts with respect to ESG practices. At a disaggregated level, the results differ between the two estimation techniques and across the CFP measures. When accounting measures are considered, the quantile analysis indicates that the Social and Governance factors are the main driving factors while the OLS results explain the ESG impact through the environmental factor. However, when the market measure Tobin's Q is employed as a proxy for CFP, the impact of each individual factors is insignificant across all quantiles, while the OLS analysis suggests the Governance factor is significant. The divergence of our results highlights the importance of acknowledging the difference between market and accounting measures, and implicitly their possible differential effect on CFP of a firm.

With respect to firms' credit ratings, the results from the probit model provide evidence of a positive impact of CSR on credit ratings in Japan at the aggregated level; on a disaggregated basis, we observe some divergence among the three pillars as there is a significant and positive effect on credit ratings based on the E and G pillars of CSR, but not the social pillar. Firms with stronger corporate governance and viewed as environmentally friendly are associated with better credit ratings, while the social pillar has less impact in the consideration of creditworthiness of issuers.

Appendix A

See Table 8.

Table 8 The constituents used in calculating the Sustainalytics' individual ESG scores

Environmental	Social	Governance
Operations	Employees	Business Ethics
Supply Chain	Supply Chain	Corporate Governance
Products and Services	Customers	Public Policy
	Community and Philanthropy	

Appendix B

See Table 9.

Table 9 Sample breakdown of number of firms by industry for each sample

	Sample 1	Sample 2
Auto Components	22	9
Automobiles	10	3
Building Products	5	4
Chemicals	30	18
Commercial Services	8	2
Construction and Engineering	12	4
Construction Materials	2	2
Consumer Durables	12	4
Consumer Services	5	3
Containers and Packaging	3	–
Diversified Metals	5	4
Electrical Equipment	9	4
Food Products	24	13
Food Retailers	11	2
Healthcare	12	1
Home Builders	6	2
Household Products	8	–
Industrial Conglomerates	3	2
Machinery	31	11
Media	10	1
Oil and Gas Producers	2	–
Paper and Forestry	2	2
Pharmaceuticals	21	3
Precious Metals	1	–
Real Estate	21	9
Refiners and Pipelines	5	4
Retailing	25	9
Semiconductors	7	4
Software and Services	19	4
Steel	9	6
Technology Hardware	30	9
Telecommunication Services	4	3
Textiles and Apparels	4	1
Traders and Distributors	9	3
Transportation	25	22
Transportation Infrastructure	3	1
Utilities	12	13
	427	182

References

Alexander, G.J., and R.A. Buchholz. 1978. Corporate social responsibility and stock market performance. *Academy of Management Journal* 21 (3): 479–486.

Ashbaugh-Skaife, H., D.W. Collins, and R. LaFord. 2006. The effects of corporate governance on firms' credit ratings. *Journal of Accounting and Economics* 42: 203–243.

Asian Bankers Association (2000). Developments of regional standards for Asian credit rating agencies—issues, challenges and strategic options. http://www.ctasc.org.tw/conference/Other%20Papers_files%5CCredit%20Ratings_files%5CCRAstudy.PDF

Attig, N., S. El Ghoul, O. Guedhami, and J. Suh. 2013. Corporate social responsibility and credit ratings. *Journal of Business Ethics* 117: 679–694.

Auer, B.R., and F. Schuhmacher. 2016. Do socially (ir)responsible investments pay? New evidence from international ESG data? *The Quarterly Review of Economics and Finance* 59: 51–62.

Aupperle, K.E., A.B. Carroll, and J.D. Hatfield. 1985. An empirical examination of the relationship between corporate social responsibility and profitability. *Academy of Management Journal* 28 (2): 446–463.

Barnett, M.L., and R.M. Solomon. 2012. Does it pay to be really good? Addressing the shape of the relationship between social and financial performance. *Strategic Management Journal* 33: 1304–1320.

Bénabou, R., and J. Tirole. 2010. Individual and corporate social responsibility. *Economica* 77 (305): 1–19.

Bhojraj, S., and P. Sengupta. 2003. Effect of corporate governance on bond ratings and yields: The role of institutional investors and outside directors. *Journal of Business* 76: 455–475.

Blume, M.E., F. Lim, and A.C. Mackinlay. 1998. The declining credit quality of U.S. corporate debt: Myth or reality. *Journal of Finance* 53: 1389–1413.

Boaventura, J.M.G., R.S.D. Silva, and R. Bandeira-de-Mello. 2012. Corporate financial performance and corporate social performance: Methodological development and the theoretical contribution of empirical studies. *Revista Contabilidade and Finanças* 23 (60): 232–245.

Bouslah, K., B. Zali, M. Turcotte, and K. Kooli. 2010. The impact of forest certification on firm financial performance in Canada and the US. *Journal of Business Ethics* 96 (4): 551–572.

Burke, L., J.M. Logsdon, W. Mitchell, M. Reiner, and D. Vogel. 1986. Corporate community involvement in the San Francisco Bay Area. *California Management Review* 28: 122–141.

Campbell, J.L. 2007. Why would corporations behave in socially responsible ways? An institutional theory of corporate social responsibility. *Academy of Management Review* 32 (3): 946–967.

Chatterji, A.K., D.I. Levine, and M.W. Toeffel. 2009. How well do social ratings actually measure corporate social responsibility? *Journal of Economics and Management Strategy* 18 (1): 125–169.

Chava, S. 2014. Environmental externalities and cost of capital. *Management Science* 60 (9): 2111–2380.

Cheng, I. H., Hong, H., and K. Shue, 2016, Do managers do good with other people's money? American Finance Association 2013 San Diego Meetings Paper; UCD and CalPERS Sustainability and Finance Symposium 2013, Fama-Miller Working Paper, Chicago Booth Research Paper No. 12-47.

Cochran, P.L., and R.A. Wood. 1984. Corporate social responsibility and financial performance. *Academy of Management Journal* 27 (1): 42–56.

Ding, G.K., C. Ferreira, and U. Wongchoti. 2016. Does it pay to be different? Relative CSR and its impact on firm value. *International Review of Financial Review* 47: 86–98.

El Ghoul, S., O. Guedhami, C.C. Kwok, and D.R. Mishra. 2011. Does corporate social responsibility affect the cost of capital? *Journal of Banking and Finance* 35 (9): 2388–2406.

Epstein, M. J. and A. Rejc-Buhovac. 2014. Making sustainability work, best practices in managing and measuring corporate social, environmental, and economic impacts. Sheffield: Greenleaf Publishing Limited

Erhemjamts, O., Q. Li, and A. Venkateswaran. 2013. Corporate social responsibility and its impact on firms' investment policy, organizational structure, and performance. *Journal of Business Ethics* 118: 395–412.

Farag, H., Q. Meng, and C. Mallin. 2015. The social, environmental and ethical performance of Chinese companies: Evidence from the Shanghai Stock Exchange. *International Review of Financial Review* 42: 53–63.

Fatemi, A., I. Fooladi, and H. Tehranian. 2015. Valuation effects of corporate social responsibility. *Journal of Banking and Finance* 59: 182–192.

Ferrell, A., H. Liang, and L. Renneboog. 2016. Socially responsible firms. *Journal of Financial Economics* 122 (3): 585–606.

Fitch Ratings. 2004. Evaluating corporate governance: The bondholders' perspective. Credit Policy Special Report.

Freeman, E. 1984. Strategic management: A stakeholder perspective. Englewood Cliffs: Prentice Hall

Friedman, M. 1970. The social responsibility of business is to increase its profits. *New York Times Magazine*, September 13, 32-33.

Garcia, A.S., W. Mendes-Da-Silva, and R.J. Orsato. 2017. Sensitive industries produce better ESG performance: Evidence from emerging markets. *Journal of Cleaner Production* 150: 135–147.

Ge, W.X., and M.Z. Liu. 2015. Corporate social responsibility and the cost of corporate bonds. *Journal of Accounting Public Policy* 34: 597–624.

Godfrey, P.C. 2005. The relationship between corporate philanthropy and shareholder wealth: A risk management perspective. *Academy of Management Review* 31: 329–346.

Goss, A., and G.S. Roberts. 2011. The impact of corporate social responsibility on the cost of bank loans. *Journal of Banking and Finance* 35: 1794–1810.

Griffin, J.J., and J.F. Mahon. 1997. The corporate social performance and corporate financial performance debate: Twenty-five years of incomparable research. *Business and Society* 36 (1): 5–31.

Jensen, M. 2001. Value maximization, stakeholder theory and the corporate objective function. *Journal of Applied Corporate Financ*e (Fall), 8–21.

Jensen, M.C., and W.H. Mecking. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3 (4): 305–360.

Jiraporn, P., N. Jiraporn, A. Boeprasert, and K. Chang. 2014. Does corporate social responsibility improve credit ratings? Evidence from geographic identification. *Financial Management* 43: 505–531.

Karpoff, J.M., J.R. Lott Jr., and E.W. Wehrly. 2005. The reputational penalties for environmental violations: Empirical evidence. *The Journal of Law and Economics* 48 (2): 653–675.

Khan, M., G. Serafeim, and A. Yoon. 2016. Corporate sustainability: First evidence on materiality. *The Accounting Review* 91: 1697–1724.

Klapper, L., and I. Love. 2004. Corporate governance, investor protection, and performance in emerging markets. *Journal of Corporate Finance* 10 (5): 703–728.

Koenker, R., and G. Bassett. 1978. Regression quantiles. *Econometrica* 46: 33–50.

Koenker, R., and K.F. Hallock. 2001. Quantile regression. *Journal of Economic Perspectives* 15 (4): 143–156.

KPMG. 2011. *KPMG international survey of corporate responsibility reporting 2011*.

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

Lamy, R., and R. Thompson. 1988. Risk premier and the pricing of primary issue bonds. *Journal of Banking and Finance* 12: 585–601.

Lins, K.V., H. Servaes, and A. Tamayo. 2017. Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *Journal of Finance* 72 (4): 1785–1824.

Lioui, A., and M. Sisto, 2017. Corporate social responsibility and the cross section of stock returns, Available at SSRN: https://ssrn.com/abstract=2730722.

Lydenberg, S., Marlin, A.T., and S. Strub, 1986. Rating America's corporate conscience: A provocative guide to the companies behind the products you buy every day, *Addison Wesley Publishing Company*.

Lys, T., J.P. Naughton, and C. Wang. 2015. Signaling through corporate accountability reporting. *Journal of Accounting and Economic* 60 (1): 56–72.

Mansi, S., M. Maxwell, and D. Miller. 2004. Does auditor quality and tenure matter to investors? Evidence from the bond market. *Journal of Accounting Research* 42: 755–793.

Margolis, J. D., Elfenbein, H.A., and J.P. Walsh, 2009. Does it pay to be good? A meta-analysis and redirection of research on the relationship between corporate social and financial performance, Available at SSRN: http://ssrn.com/abstract=1866731

Mattingly, J.E., and S.L. Berman. 2006. Measurement of corporate social action: discovering taxonomy in Kinder Lydenberg Domini Ratings Data. *Business and Society* 45: 20–46.

McWilliams, A., and D. Siegel. 2000. Corporate social responsibility and financial performance. *Strategic Management Journal* 21 (5): 603–609.

Menz, K.M. 2010. Corporate social responsibility: Is it rewarded by the corporate bond market? A critical note. *Journal of Business Ethics* 96: 117–134.

Nguyen, P.A., Kecskes, A., and S. Mansi. 2017. Does corporate social responsibility create shareholder value? The importance of long-term investors. *Journal of Banking and Finance* pp. 1–21.

Oikonomou, I., K. Brooks, and S. Pavelin. 2014. The effects of corporate social performance on the cost of corporate debt and credit ratings. *The Financial Review* 49: 49–75.

Orlitzky, M., and J.D. Benjamin. 2001. Corporate social performance and firm risk: A meta-analytic review. *Business and Society* 40: 369–396.

Ottawa, B. 2018. Not considering ESG could mean failing fiduciary duty: Compenswiss. *IPE,* Available at https://www.ipe.com/news/esg/not-considering-esg-could-mean-failing-fiduciary-duty-compenswiss/www.ipe.com/news/esg/not-considering-esg-could-mean-failing-fiduciary-duty-compenswiss/10025298.fullarticle.

Preston, L.E., and D.P. O'bannon. 1997. The corporate social-financial performance relationship: a typology and analysis. *Business and Society* 36 (4): 419–429.

Principles for Responsible Investment. 2016. Global guide to responsible investment regulation. PRI *Association.*

Rodgers, W., H.L. Choy, and A. Guiral. 2013. Do investors value a firm's commitment to social activities? *Journal of Business Ethics* 114: 607–623.

Rust, R., K. Lemon, and V. Zeithaml. 2004. Return on marketing: Using customer equity to focus marketing strategy. *Journal of Marketing* 68: 109–127.

Sahut, J.M., and H. Pasquini-Descomps. 2015. ESG impact on market performance of firms: International evidence. *International Management* 19 (2): 40–63.

Sharfman, M.P., and C.S. Fernando. 2008. Environmental risk management and the cost of capital. *Strategic Management Journal* 29: 569–592.

Spencer, B.A., and G.S. Taylor. 1987. A within and between analysis of the relationship between corporate social-responsibility and financial performance. *Akron Business and Economic Review* 18 (3): 7–18.

Teoh, S.H., I. Welch, and C.P. Wazzan. 1999. The effect of socially activist investment policies on the financial markets: Evidence from the South African boycott. *The Journal of Business* 72 (1): 35–89.

Tsoutsoura, M. 2004. Corporate social responsibility and financial performance, *Center for Responsible Business*.

Ullmann, A. 1985. Data in search of a theory: A critical examination of the relationships among social performance, social disclosure, and economic performance of U.S. firms. *Academy of Management Review* 10: 540–557.

Vance, S.C. 1975. Are socially responsible corporations good investment risks? *Management Review* 64 (8): 19–24.

Waddock, S.A., and S.B. Graves. 1997. The corporate social performance—financial performance link. *Strategic Management Journal* 18 (4): 303–319.

Wolfe, R., and K. Aupperle. 1991. Introduction to corporate social performance: methods for evaluating an elusive construct. In J. E. Post (Ed.), *Research in Corporate Social Performance and Policy* 12, 265-68.

Wu, M.L. 2006. Corporate social performance, corporate financial performance, and firm size: A meta-analysis. *Journal of American Academy of Business* 8 (1): 163–171.

Xie, X.M., Y.Y. Jia, X.H. Meng, and C. Li. 2017. Corporate social responsibility, customer satisfaction, and financial performance: The moderating effect of the institutional environment in two transition economies. *Journal of Cleaner Production* 150: 26–39.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Frank J. Fabozzi is a Professor of Finance at EDHEC Business School and a Senior Scientific Adviser at EDHEC-Risk Institute. He is a trustee of the BlackRock fixed-income fund complex. He is the editor of *The Journal of Portfolio Management* and cofounder and coeditor of *The Journal of Financial Data Science*. He is the CFA Institute's 2007 recipient of the C. Stewart Sheppard Award and the CFA Institute's 2015 recipient of the James R. Vertin Award. He was inducted into the Fixed Income Analysts Society Hall of Fame in November 2002. Fabozzi received his BA and MA in economics in 1970 from The City College of New York where he was elected to Phi Beta Kappa and a PhD in economics in 1972 from the City University of New York. He has earned the designations of CFA and CPA.

Peck Wah Ng is currently a Director in MUFG Bank, Singapore, and has over 20 years of experience in banking, specifically in risk management. She started her career with the Monetary Authority of Singapore and has held senior positions in other financial institutions including BNP Paribas, Deutsche Bank and China Citic Bank International. Ng holds a PhD in Finance from EDHEC Business School. She received her Bachelor of Accountancy and MBA in Banking and Finance from Nanyang Technological University and completed the International Management Program at New York University's Stern School of Business. She has been conferred both the CFA and CA (Singapore) designations.

Diana E. Tunaru is a Lecturer in Finance at Kent Business School, University of Kent, UK. She holds a PhD in financial economics (2017) from University of Westminster, London. As an early career researcher, her areas of interest cover continuous-time modeling in general, modeling of the term structure of interest rates in particular, volatility spillovers, econometric analysis and more recently responsible investing/green finance.

Journal of Asset Management (2021) 22:96–109
https://doi.org/10.1057/s41260-020-00192-z

ORIGINAL ARTICLE

Green bonds: shades of green and brown

Moritz Immel[1] · Britta Hachenberg[2] · Florian Kiesel[3] · Dirk Schiereck[1]

Revised: 13 October 2020 / Accepted: 14 October 2020 / Published online: 29 October 2020

Abstract
We analyse the existence of a green bond premium and find a negative premium of 8 to 14 basis points. We are further interested in the influence of ESG ratings on green bonds to determine if investors differentiate between the shade of green. Examining a unique dataset of green bonds, we find a statistically significant influence of ESG ratings on bond spreads. A one-point increase in the weighted average ESG score leads to a decrease in the spread of 6 to 13 basis points. Interestingly, the results are not driven by the environmental friendliness of the green bond issuer, but through the company's governance.

Keywords Green bonds · ESG ratings · Green bond premium · Governance

JEL Classification G12 · M14 · Q50

Introduction

A few years ago, green bonds were mainly issued by government-related or supranational development banks. Green bonds are securities, whose proceeds are used to support climate-related or environmental projects. During the last years, more and more corporates have started to issue green bonds, as well, followed by governments who discovered this asset class. While there have been many constructive developments in this segment, not all have been perceived positively. The accusation of greenwashing is omnipresent, and the "true greenness" of green bonds is regularly discussed (see, e.g., Laufer 2003; Wu et al. 2020). One reason for this uncertainty is that green bonds are usually issued with a credit rating but provide additionally environmental, social and governance (ESG) ratings only on a voluntary basis. We focus on this ESG rating information and examine the influence of ESG aspects on the pricing of green bonds, specifically the bond spread.

The potential existence of a (negative) green bond premium has been analysed by manifold studies in the past (e.g., Hachenberg and Schiereck 2018; Zerbib 2019), but the results are mixed. The green bond premium is defined as the incremental yield investors receive for holding a green bond over its equivalent non-green counterpart. The non-green counterpart is often referred to as a conventional or brown bond and implies no specific use of proceeds. Hence, the bonds we analyse are exclusively green or brown. Therefore, we define the bonds we analyse as either green or brown. In a first step, we test if a green bond premium can be found using a similar approach as Preclaw and Bakshi (2015). In a second step, we analyse the influence of ESG ratings on green bond pricing. We define the directional effect as dependence of the green bond premium on the existence of the ESG rating. ESG ratings are based on issuer level and graded on a scale. If the company's shade of green matters for the pricing of green bonds, the (negative) premium is expected to be larger for higher ESG ratings. We define this as the magnitude effect. As shades of green (or brown) are hardly the subject of investigations in green bond studies, we contribute to the existing literature on green bonds and fill a niche examining the greenness of the bond. Finally, as ESG ratings are a composition of E, S and G criteria, it is obvious to test which of these criteria are the main drivers of the green bond premium. We define this as the composition effect.

The rest of the paper is structured as follows. The next section provides a literature review and develops our hypotheses. "Data and methodology" section presents the data and

✉ Florian Kiesel
florian.kiesel@grenoble-em.com

1 Technical University of Darmstadt, 64289 Darmstadt, Germany

2 Technische Hochschule Köln, 50678 Cologne, Germany

3 Grenoble Ecole de Management, 38000 Grenoble, France

methodology used. "Empirical results" section documents the empirical results, and "Conclusion" section concludes the paper and outlines possible areas of future research.

Sample literature review and hypothesis development

ESG has been the subject of research for many years now, however, not necessarily using the term ESG. Research on topics such as corporate social responsibility (CSR) or corporate environmental responsibility (CER) goes back more than 60 years.

Many studies were published since the 1990s, most of them focussing on empirical relationships between corporate social/environmental responsibility and corporate financial performance (CFP), which might be due to improved data availability (Capelle-Blancard and Monjon 2012). The majority of papers find a positive relationship between ESG and CFP. The relationship is measured, among others, through market-stock prices or accounting measures, e.g., return on assets (Schiereck et al. 2019).

Since the pricing of bonds is more complicated and technical due to the huge variety of bonds (different coupon type, maturity, payment rank, callability, etc., see, e.g., Maul and Schiereck 2017), we specifically review the literature regarding the pricing of bonds. By analysing 4260 bonds issued between 1992 and 2009, Ge and Liu (2015) find that CSR performance is associated with better credit ratings and lower yield spreads in new corporate bond issues.

The majority of empirical studies investigating green bonds focus on the (negative) green bond premium. Panel A of Table 1 summarizes the most relevant literature analysing the green bond premium. Zerbib (2019) compares 110 senior fixed-rate green bonds with their equivalent synthetic conventional bonds. He finds a significant negative green bond premium of approximately 2 bps. The effect is especially visible for financials as well as for lower rated bonds. For government-related issuers, the effect is not visible on a AAA-rating base, but a small negative premium can be found for AA-rated issuers.

Hachenberg and Schiereck (2018) also use a matching approach in order to avoid heterogeneity among bonds. They find that A-rated bonds show a significant negative premium of 3.88 bps. Green bonds with different ratings also trade tighter than their corresponding brown bonds but without statistical significance. The study shows that the only significant factors for the "Greenium" are industries and the existence of an issuer's ESG rating. Contrary to Zerbib (2019), they find that green bonds from government-related companies trade wider than comparable brown bonds.

Bachelet et al. (2019) show that green bonds have higher yields coupled with higher liquidity and are also less volatile than their brown correspondents. By investigating the issuer breakdown further, the authors determine a difference between institutional and private issuers. Green bonds from institutions have a negative yield premium and are far more liquid than the matching brown bonds. Green bonds from private issuers on the other hand have a positive premium and do not differ too much in liquidity compared to brown ones. Moreover, if the private issuer has no third-party verification, the premium is significantly higher. Bachelet et al. (2019) conclude that institutional investors are able to attract large institutional investors as these issuers have transparency and information rules that lower information asymmetries. Private issuers can get a similar reputational effect when they obtain a green verification from a third party.

Partridge and Medda (2018) run a yield curve analysis on a selection of US green-labelled municipal bonds, issued at the same time as brown bonds and by the same issuers. Additionally, they use a pair-wise analysis to identify potential yield differentials between bonds that are identical except for the green label (same issuer, maturity, etc.). For both analyses, they find a (negative) green premium in the primary as well as in the secondary market. Contrary, Karpf and Antoine (2017) find that the green characteristic of a green-labelled municipal bond gets penalized through the market, as green bonds trade at higher yields than expected by their credit profiles.

Preclaw and Bakshi (2015) investigate that a premium is being paid by investors in order to acquire green bonds. They calculate a spread difference of approximately 20 bps between green bonds and ordinary (brown) bonds. They quantify this green bond premium using a regression that decomposes the OAS-spread into common risk factors like credit rating or spread duration and a dummy variable for green bonds. Preclaw and Bakshi (2015) use a global credit index including corporates as well as government-related issuers, which provides the best overlap of constituents with the global green bond index.

Using a propensity score matching approach, Gianfrate and Peri (2019) find that green bonds are cheaper to issue than ordinary bonds. This effect holds even after accounting for green certification costs and is larger for corporates. This finding is contrary to Kapraun and Scheins (2019) who find that green bonds are generally traded at a higher bond premium except for green bonds issued by governments and supranational institutions.

Summarizing the discussed literature, there is no clear consensus about a (negative) green bond premium. Therefore, the first question we aim to answer is whether a green bond premium really exists.

Should this green bond premium exist and confirmed to be negative, we must address the concern posed by a potential greenwashing effect, i.e., that green bond issuers may be attempting to present a misleading impression of

Table 1 Empirical findings related to this study

Study	Green bonds (GBP)	Scope	Primary/Secondary market	Number of bonds	Time period	Method	Yield premium	Main findings
Panel A: Empirical studies related to green bonds								
Zerbib (2019)	Yes	Global	Secondary	110 green bonds/110 synthetic bonds	July 2013–December 2017	Matching + 2 step regression procedure	−2 bps	The negative premium is more pronounced for financial and low-rated bonds. The results emphasize the low impact of investors' pro-environmental preferences on bond prices.
Hachenberg and Schiereck (2018)	Yes	Global	Secondary	63 green bonds/126 brown bonds	October 2015–March. 2016	Matching + panel regression	−1 bp	Significant are neither maturity, nor volume or currency, but rather industries, namely government-related and financial issuers, as well as the existence of an ESG issuer rating
Bachelet et al. (2019)	Yes	Global	Secondary	89 green bonds + 89 brown bonds	January 2013–Dez. 2017	Matching	2.06–5.9 bps	Green bonds have higher yields. Premium for non-certified green bonds are higher than for those certified. Institutional green bonds have a negative premium.
Partridge and Medda (2018)	N.A.	US green-labelled muni bonds	Primary and secondary	1215 (548 green and 667 brown)	June 2013–January 2018	Yield curve analysis + matching	−5 bps in the secondary market	They find that there is a growing trend towards green premium in both the primary and secondary markets in both the series trend analysis and in the pair-wise analysis

Table 1 (continued)

Study	Green bonds (GBP)	Scope	Primary/Secondary market	Number of bonds	Time period	Method	Yield premium	Main findings
Karpf and Antoine (2017)	N.A.	US green-labelled muni bonds	Secondary	1,880 green bonds	N.A.	Yield curve analysis	Positive premium	The market valuates green bonds in a less favourable manner than their brown counterparts. If green bonds had the same coefficients as brown bonds or the pooled sample, their expected mean return would be lower - > the market penalizes green bonds to a higher degree than brown bonds
Preclaw and Bakshi (2015)	Yes	Euro and US	Secondary	N.A.	March 2014–August 2015	OLS regression	− 17 bps	Negative premium
Gianfrate and Peri (2019)	N.A.	Europe	Primary and secondary	121 green bonds	2013–2017	Propensity score matching	Negative premium	Magnitude depending on secondary or primary market and on corporate and non-corporate issuers.
Kapraun and Scheins (2019)	Yes	Global	Primary and secondary	1,500 green bonds	N.A.	OLS regression	− 18 bps	Especially for corporate green bonds, a green label, a third-party verification or a listing on a dedicated green bond exchange is decisive in order to be seen as a green bond and to generate the premium.

Table 1 (continued)

Study	Purpose of investigation	Main findings
Panel B: Studies analysing the impact of ESG ratings		
Polbennikov et al. (2016)	ESG rating influence on corporate bond spreads	Bonds with high composite ESG ratings have slightly lower spreads, all else being equal. They also find that bonds with high ESG ratings have modestly outperformed their lower rated peers when controlling for various risk exposures
Menz (2010)	CSR-influence on corporate bond spreads	CSR has apparently not yet been incorporated into the pricing of corporate bonds
Stellner et al. (2015)	CSR-influence on credit rating and z-spread	Only weak evidence of unconditional benefits from CSR investments on the z-spread. But in countries with above average ESG rating better CSP performance is rewarded with a better rating and a lower spread
Gatti and Florio (2018)	Influence of Green Bond Principles and second party review on green bond spreads	Green bond issues with a second party opinion have a lower spread than those without
Barnett and Salomon (2012)	Relationship between CSR and CFP	Low social responsibility comes with a higher CFP than moderate social responsibility but with a lower CFP than high social responsibility
Trumpp and Guenther (2017)	Relationship between CEP and CFP	Companies with low CEP have a negative relationship to CFP, while companies with a very high CEP are positively related.
Nollet, Filis and Mitrokostas (2016)	Relationship between CSR and CFP in the S&P 500	Applying a linear model leads to a negative relationship between CSR and CFP. But applying a nonlinear model they find a u-shaped relationship which implies that there is a threshold amount of investments going into CSR after which the engagement will show positive effects with regards to the financial performance

their activities and the lack of consistency in audit standards. We have a number of tools at our disposal including the existence of an ESG rating or an external certification and we focus our efforts on the former. Testing the dependence of a green bond premium on the existence of an ESG rating, called the directional effect, is the subject of studies from, among others, Polbennikov et al. (2016), Menz (2010) and Stellner, Klein and Zwergel (2015). Polbennikov et al. (2016) measure slightly lower spreads for corporates with higher ESG ratings. This strand of literature is summarized in Panel B of Table 1. Menz (2010) cannot confirm that CSP/CEP/ESG is related to lower financing costs. His study shows that firms that are labelled socially responsible have a higher risk premium than non-socially responsible companies.

Stellner et al. (2015) empirically observe the influence of CSR on the credit rating and the z-spread. They find only weak evidence of unconditional benefits from CSR investments on the z-spread. But by examining more closely the influence of the issuer's country, they observe that in countries with above average ESG ratings better CSP performance is rewarded with a better rating and a lower spread. Additionally, they find that it is beneficial for companies to have the same relative ESG rating as the country (above average or below average).

Gatti and Florio (2018) investigate the role of the Green Bond Principles and a second party review on green bond spreads. Using a sample of green bonds issued between 2007 and 2015, they find that with the introduction of the Green Bond Principles in 2014 issues with low credit ratings were also able to enter the market.

The certification of green bonds is a field of analysis of Bachelet et al. (2019), as well. They extend their study and divide their sample of private issuers in certified and non-certified green bonds. For non-certified issues, they detect a positive premium. They conclude that green bonds can have a negative premium (lower financing costs) under the premise of trust which is either generated through being an institution or through green verification. Missing reputation or certification will lead to higher financing costs due to the investors' concerns about greenwashing.

Kapraun and Scheins (2019) show similar results analysing the pricing in both primary and secondary markets of a sample of more than 1,500 green bonds. For both markets, only certain green bonds show lower yields (i.e., negative green premium) in comparison with their brown counterparts. This applies in particular to issues of government and supranational entities, butalso corporate issues when they issue at large size. The latter is in contrast to Zerbib (2019). Especially for corporate green bonds, a green label, a third-party verification or a listing on a dedicated green bond exchange is decisive in order to be seen as a green bond and to generate the negative premium.

Hachenberg and Schiereck (2018) find that having an ESG rating reduces the negative premium (green bonds are priced less tight than brown bonds). This might seem surprising at first thought, but they argue that this might be due to the fact that ESG-dedicated investors do not necessarily need to pick a green bond where the issuer has an ESG rating in order to conform with their ESG investment policy. The ESG rating might allow the investor to simply purchase the ordinary bond. By extending this research, we analyse further determinants of green bond pricing. Hachenberg and Schiereck (2018) compare green bonds with ordinary (brown) bonds. We, on the other hand, compare green bonds from various issuers with distinct characteristics with each other.

Based on the literature, we develop our first hypothesis. The assumption of Hypothesis 1 is that a green bond with ESG rating has a lower spread due to reduced uncertainty about the bonds shade of green. This potential divergence is called the directional effect and matches Gatti and Florio (2018) and Bachelet et al. (2019) who show that verification is associated with lower spreads. Even though an ESG rating alone is not a valid verification, it reduces the information asymmetry between issuer and investor regarding the greenness of the bond and potential greenwashing.

Hypothesis 1 Existence of an ESG rating leads to higher credibility!

Next, we will look at the specific characteristics of ESG and the magnitude effect. Recent research emphasizes that the relationship between ESG and CFP does not necessarily have to be linear. The idea behind this is that "too much of a good thing" can have negative consequences (Pierce and Aguinis 2013). Barnett and Salomon (2012) show that the relationship between CSR and CFP is u-shaped. Meaning that low social responsibility comes with a higher CFP than moderate social responsibility, but with a lower CFP than high social responsibility. Trumpp and Guenther (2017) find similar results investigating the relationship between CFP and corporate environmental performance (CEP). Companies with low CEP have a negative relationship to CFP while companies with a high CEP are positively related. They call this the "too little of a good thing" effect.

A negative green premium might be due to high demand for this new asset class, which fits quite well into the current political situation of growing environmental concern. Therefore, it is necessary to review the influence of CSP/CEP/ESG on bonds in general. We are especially interested in the influence on green bonds, as we aim to find out if green investors actually care about the shade of the green bond or if only the label counts.

We analyse the influence of ESG ratings on green bond pricing to determine if the greenness of the issuer matters for its pricing. A better ESG rating should lead to lower spreads as already found in some literature and corresponding with the CSP/CEP-CFP research (e.g., Polbennikov et al. 2016; Zerbib 2019).

Hypothesis 2 The better the ESG rating, the lower the spread!

Nollet, Filis and Mitrokostas (2016) examine the CSR and CFP relationship by using the S&P 500 universe and taking the Bloomberg ESG disclosure score as a proxy for CSP. Applying a linear model, they show a negative relationship between CSR and CFP. But by applying a nonlinear model, they find a u-shaped relationship. This implies a threshold amount of investments going into CSR, before the engagement will show positive effects with regards to the financial performance. When splitting up the ESG score into E, S and G, they show that the governance aspect is the main driver for translating CSR into CFP.

Finally, we separately analyse the impact of E, S and G on green bond pricing in order to draw conclusions on the composition effect. We expect that in particular the E- and G-score should have a significant effect on the bond spread as a negative green bond premium indicates that investors accept getting paid less through green assets. Therefore, a green bond issued with a better ESG rating, in particular a better E-rating, should have a lower spread than a bond issued with a lower ESG rating. In particular, we expect the E-score to have an influence (negative correlation to spread), as our objects of investigation are green assets. We assume that the social score is less important while the governance score should also be relevant as a low score would indicate low issuer trustworthiness. Since trust is important regarding the use of proceeds, the G-score is expected to be negatively correlated to the spread.

Hypothesis 3 For green bonds, environmental criteria dominate social and governance criteria!

Data and methodology

We use three different datasets for our analyses. Dataset 1, the Bloomberg Barclays Global Aggregate Index which will be used to examine whether a (negative) green bond premium exists. Dataset 2, a unique screening of the fixed universe to produce a sample of 466 bonds. Details of this screening process are provided below. This dataset is used to determine whether and how the issuers' ESG rating affects the respective green bond spreads. Finally, dataset 3, the ICE Bank of America Merrill Lynch Green Bond Index which is used to control the robustness of the results and to validate the screening rules applied to dataset 2. We will henceforth refer to dataset 1, dataset 2 and dataset 3 as "the global aggregate index", "the custom universe" and "the green bond index".

We focus on the results of the custom green universe and use the green bond index and the green component of the aggregate index for validation purposes. The aggregate index includes green bonds and therefore these (green) bonds can be used as an additional database for validation of hypotheses 1 and 2. We apply MSCI ESG ratings to verify the impact of ESG ratings on bond spreads. There is a growing stream of literature that documents the divergence of ESG ratings (see, e.g., Berg et al. 2019; Chatterji et al. 2009, 2016; Semenova and Hassel 2015; Dorfleitner et al. 2015). The recent study of Berg et al. (2019) shows that the correlation between five ESG raters ranges between 0.42 and 0.73. This divergence can be mainly explained by measurement divergence, and the number of categories the ESG provider is using. In their findings, they show that while MSCI (previously KLD but was acquired by MSCI) nearly needs 25 categories to regress the rating, the other four rating providers need significantly less categories to explain their ratings. Their findings show how the divergence between the rating providers might explain the difference in the rating. However, the results also indicate that MSCI has the best overall ESG score and that could explain why it is the most used one in academic literature (Berg et al. 2019) to verify the impact of ESG ratings on bond spreads.

Data are as at October 31st, 2019. By using secondary bond spreads instead of primary spreads, we reduce the influence of possible macroeconomic influences. We obtain all bond control variables from Bloomberg.

As the first green bond was issued in 2007, we reduce the debt universe to bonds issued between January 1st, 2007 and October 31st, 2019. Next, we filter for green bonds. For the observed time period, 2456 issues are green labelled. As this sample also includes loans, we excluded them. Thereafter, we implement a size threshold of $100 million issue volume. This step reduces the sample down to 1077 bonds. In order to create a homogenous dataset, we look at bonds with maturity type "At Maturity" and "Callable" only. This reduces the sample size by 64 observations. Further homogenization is reached by excluding floating rate bonds. We also exclude bonds without a credit rating.[1] This reduces the sample by another 328 bonds. Adjusting for double counting through RegS and 144a issues as well as Tap issues the sample is left with 493 bonds. For further homogeneity 18 bonds whose coupon types are not fixed are removed. As a last step, 9 bonds are excluded due to missing data.

Table 2 Sample selection process

	Number of securities
Initial sample	2456
Less loans	– 109
Less bonds with amount issued < 100 mio $	– 1270
Less bonds with maturity type different from "at maturity" or "callable"	– 64
Less floating bonds	– 82
Less bonds with no credit rating	– 328
Less bonds with series "RegS"	– 57
Less bonds that are taps	– 53
Less bonds with coupon type different from "fixed"	– 18
Less bonds that miss necessary data (e.g. I-spread)	– 9
Final sample	466

The sample selection process reduces the number of green bonds from 2456 down to a final sample of 466 bonds. Table 2 summarizes the sample selection.

To test hypotheses 1 to 3, we need to further adjust the dataset. For Hypothesis 1, we use the final sample, respectively, the green bond index and the green part of the global aggregate index[2] for validation. For hypotheses 2 and 3, only bonds from issuers with an ESG rating can be used. Therefore, we match bonds with the issuers MSCI ESG rating. If a subsidiary who is not rated issued the bond, we used the ultimate parent's ESG rating. This method of matching ratings and bonds is rather problematic for government-related issuers, as the ultimate parent is (ultimately) the government. The matching for government-related issuers can therefore be inaccurate. To avoid this problem, we use two different datasets for each regression. The first dataset is our full dataset and the other dataset incluses coporate bonds only. When discussing the results, we will focus on corporate green bonds.

In order to determine the influence of different variables on bond spreads, especially the ESG rating, an ordinary least squares (OLS) regression is applied as followed:

$$\begin{aligned} Y_i = {} & \beta_0 + \beta_1 \text{ESG}_i + \beta_2 \text{Credit Rating}_i + \beta_3 \ln\left(\text{Amount Issued in \$}\right)_i \\ & + \beta_4 \ln\left(\text{Amount outstanding in \$}\right)_i + \beta_5 \ln\left(\text{Time to maturity in years}\right)_i \\ & + \beta_6 \text{Callable}_i + \beta_7 \text{USD}_i + \left(\beta_8 \text{Governm. related}_i\right) \\ & + \beta_9 \text{China}_i + \beta_{10} \text{Payment Rank}_i + u_i \end{aligned} \tag{1}$$

[1] We forgo to use a rating of a comparable bond. Credit ratings lower the information asymmetry between the issuer and the investors, and this might lead to a bias of the results in using comparable bonds.

[2] Floaters are not included in the green bond index as well as the global aggregate index.

Table 3 Variable definitions

Variable	Description
ESG related variables	
ESG rating	Dummy variable which takes value 1 if the issuer has a MSCI ESG rating, 0 otherwise
E-score	Environmental pillar of ESG Score from 0 (worst) to 10 (best)
S-score	Social pillar of ESG Score from 0 (worst) to 10 (best)
G-score	Governance pillar of ESG Score from 0 (worst) to 10 (best)
Weighted average ESG score	Combined ESG score from 0 (worst) to 10 (best)
Control variables	
Credit rating	Bloomberg composite rating (expanded if not available with S&P or Moody's rating.) AAA equals 1, AA + equals 2 etc
Amount issued	Issue size in USD
Amount outstanding UP	Amount outstanding (all bonds) of the ultimate parent in USD.
Time to maturity	Remaining time to maturity in years measured from October, 31, 2019
Callable	Dummy variable which takes value 1 if the bond is callable, 0 otherwise
USD	Dummy variable which takes value 1 if the bond is denominated in USD, 0 otherwise
Governm. related	Dummy variable which takes value 1 if the bond is issued by a government-related issuer, 0 otherwise
Green bond	Dummy variable which takes value 1 if the bond has a Green Instrument flag, 0 otherwise
China	Dummy variable which takes value 1 if the issuers country of risk is China, 0 otherwise
Payment rank	Normalized payment rank of the bond where 1 equals 1st lien Secured, 2 equals Secured, 3 equals Sr unsecured, 4 equals Subordinated, and 5 equals Jr Subordinated

Variables are described in Table 3. Slight changes to the base models are necessary depending on the hypothesis tested or the data used.[3]

Empirical results

Does a (negative) green bond premium exist?

To determine whether there is a statistically significant green bond premium, we follow the approach of Preclaw and Bakshi (2015). We use the global aggregate index and introduce a green bond dummy variable, which is one if the bond is green and zero otherwise. Table 4 shows the regression results. We control for collinearity of the variables by analysing the variance inflation factors (VIFs). The average VIFs are low and around 1.54 for the dataset including all bonds and 1.45 for the dataset focusing on corporates; hence, we may assume there is no collinearity of the variables.

The *Green bond* dummy variable is statistically significant and negative, indicating a negative green bond premium of 8 to 14 bps. These results suggest that investors are willing to receive a lower yield in order to buy green. Thus, we confirm a (negative) green bond premium. The findings are in line with Zerbib (2019).

Does having an ESG rating lower the spread?

Next, we test the directional effect of an ESG rating on the spreads of green bonds. To determine whether the ESG rating has a positive impact on spreads of green bonds, we analyse if a missing ESG rating leads to a higher spread. The ESG variable is a dummy variable with the value of one if the issuer has an ESG rating (from MSCI) and zero otherwise.

Table 5 shows the results of the OLS regression. It includes six different regressions, but our main focus is on the regressions that include corporates only.

Analysing the influence of *ESG Rating*, we find that all three corporates only regressions (Models 1, 3, 5) show a negative relation of ESG rating and spread. Model 1, however, shows no statistically significant effect of the *ESG* variable, but the regressions of the green bond index (Model 3) and the green part of the global aggregate index (Model 5) are both highly statistically significant. Depending on the estimated model, having an *ESG rating* lowers the spread by 9 to 19 bps. The insignificant results of Model 1 could be due to a dominance of green bonds that have an *ESG rating*.

A deterioration in *Credit Rating* of 1 step (e.g., AA to AA-) leads to a 28.66 bps higher spread (Model 1 of Table 5). The influence of the *Credit Rating* is positive and statistically significant for each regression. The positive coefficient is not surprising, as a lower rating indicates more risk. The *Amount Issued* is statistically significant at the 5% level for the green component of the global aggregate index

[3] Depending on the model, one or three ESG variables are used.

Table 4 Existence of a green bond premium

	Green bonds characteristic only		Green bond characteristic and ESG Yes/No		Green bond characteristic and E-,S- and G-Score	
	Corporates	All	Corporates	All	Corporates	All
Model	1	2	3	4	5	6
Green bond	−9.704*** (1.909)	−8.030*** (1.630)	−12.777*** (1.883)	−11.061*** (1.618)	−13.713*** (2.158)	−12.215*** (2.066)
ESG			−22.140*** (0.929)	−20.829*** (0.838)		
E-score					−2.495*** (0.204)	−2.467*** (0.201)
S-score					−1.407*** (0.267)	−1.281*** (0.267)
G-score					−1.543*** (0.310)	−1.205*** (0.303)
Credit rating	13.176*** (0.174)	12.206*** (0.155)	13.410*** (0.174)	12.412*** (0.155)	11.775*** (0.204)	11.248*** (0.197)
Amount issued	−4.621*** (0.617)	−5.220*** (0.422)	−2.710*** (0.611)	−2.794*** (0.432)	−2.130*** (0.624)	−1.858*** (0.483)
Time to maturity	41.111*** (0.463)	38.214*** (0.428)	40.607*** (0.456)	37.891*** (0.420)	43.980*** (0.506)	41.341*** (0.482)
Callable	−23.596*** (0.802)	−20.413*** (0.784)	−18.791*** (0.777)	−16.185*** (0.756)	−16.911*** (0.825)	−15.484*** (0.825)
USD	37.409*** (0.733)	39.404*** (0.684)	36.931*** (0.712)	38.573*** (0.664)	27.166*** (0.772)	29.210*** (0.746)
Payment rank	−3.193 (0.521)	−1.954*** (0.494)	−2.004*** (0.525)	−0.644 (0.498)	−2.743*** (0.685)	−2.149*** (0.680)
Intercept	16.237 (12.536)	33.006*** (8.626)	−12.921 (12.330)	−7.771 (8.756)	−3.119 (13.065)	−7.099 (10.253)
N	14,170	16,046	14,170	16,046	10,705	11,543
Adjusted R^2	0.65	0.67	0.67	0.68	0.68	0.69

Standard errors are reported in parentheses

$^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$

(Model 5). This result holds when including government-related bonds (Model 6). This relationship is in line with the previous literature (Kapraun and Scheins 2019). For the green bond index (Model 1 and 2), the coefficients are not statistically significant, which might be due to the issue size threshold that was implemented. A required minimum issue size might lower the influence of issue size on the spread. The *Amount Outstanding* variable shows only weak significance for the green dataset including all bonds (Model 2).

The relationship between *Time to Maturity* and spread is positive, i.e., a longer time to maturity leads to a higher spread. For all regressions, this influence is statistically significant. The coefficient is in the range of 11–26. The loading factors for the variable in the regression analysis suggest, that, all else being equal, an increase of the time to maturity by 1 year is associated with a higher spread of between 6.808 bps and 22.737 bps. The direction of the influence is as expected.

The *Callability* of a green bond leads to a 44.7 bps lower spread (Model 1) and is statistically significant for all but one regression (Model 4). This is contrary to previous research (e.g., Kuhn et al. 2018) but could be due to the effect of the low-interest phase. If an increase in interest rates is expected rather than a further decrease the *Callability* of a bond is no disadvantage for an investor. However, the results should be interpreted carefully as the callability also depends on the accepted likelihood of being called. Alternatively, the call option has the feature of reducing the maturity and hence could lead to lower spreads if compared to a higher maturity non-callable bond.[4]

For the *USD* dummy variable, we find that USD denominated bonds are significantly wider in a range of 46 to 74

[4] We thank the anonymous reviewer for this alternative explanation.

Table 5 Regression results Hypothesis 1

	Custom universe		Green bond index		Global aggregate index	
	Corporates	All	Corporates	All	Corporates	All
Model	1	2	3	4	5	6
ESG rating	−9.512 (16.177)	−36.412*** (9.611)	−19.179*** (5.952)	−18.622*** (4.623)	−12.636*** (3.869)	−12.531*** (3.677)
Credit rating	28.655*** (3.614)	21.601*** (2.370)	12.252*** (1.320)	10.982*** (0.916)	11.837*** (0.818)	11.371*** (0.755)
Amount issued	3.423 (16.695)	−7.440 (5.818)	−5.148 (6.714)	−6.203** (2.675)	−13.288** (5.134)	−14.965*** (3.724)
Amount outstanding	−2.389 (2.578)	−4.994* (2.596)				
Time to maturity	17.100* (9.467)	11.743** (4.788)	26.311*** (3.262)	18.922*** (1.988)	26.440*** (2.956)	24.034*** (2.186)
Callable	−44.705** (19.885)	−36.799** (17.218)	−7.787* (4.549)	0.207 (4.538)	−19.951*** (4.178)	−16.344*** (3.990)
USD	74.424*** (15.753)	71.911*** (9.806)	56.641*** (3.924)	46.427*** (2.976)	51.596*** (3.387)	47.056*** (2.840)
Governm. related		−7.062 (8.724)		−14.090*** (4.067)		−10.787*** (3.324)
China	295.445** (137.771)	63.727* (36.236)	−2.724 (8.898)	0.082 (8.126)		
Payment rank	−76.575*** (26.777)	−59.191*** (20.007)	4.477 (9.351)	1.499 (9.220)	−10.894*** (3.488)	−10.977*** (3.540)
Intercept	61.010 (348.662)	378.628 (124.235)	34.455 (126.748)	84.503* (46.210)	247.955** (101.488)	289.409*** (72.886)
N	218	466	301	491	323	407
Adjusted R^2	0.67	0.60	0.61	0.68	0.74	0.75

Standard errors are reported in parentheses

$^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$

bps.[5] It can be concluded that other currencies, such as Euro denominated green bonds, price tighter than USD denominated green bonds.

The variable *Government related* is only relevant for the regressions using the full sample. All three regressions (Models 2, 4, 6) show that government-related issuers receive a lower spread of 7 to 14 bps. However, only Models 4 and 6 show statistical significance for this variable. The dummy variable *China* only shows significance for the custom dataset (Model 1 and 2). The "corporates only" (Model 1) selection indicates that a bond from China has a 295 bps higher spread, while the whole sample shows an additional spread of only 64 bps. In general, the higher spread related to *China* seems comprehensible, as characteristics of green bonds in China differ from those of other markets. The China Green Bond Market Report (2019) points out that a high percentage of bonds from China and labelled green does comply with the internationally recognized definition of a green bond by the Climate Bonds Initiative. On the other hand, there are also green-labelled bonds that only comply with China's domestic definitions. Another takeaway is that Chinese companies that are not government related seem to be regarded as much riskier, which might be due to the government's strong influence on the economy.

The results of the influence of the *Payment Rank* are mixed. The green bond index (Model 3 and 4) indicates a positive relationship (1 = 1st lien Secured … 5 = Jr. Subordinated) between spread and payment rank, but these results are not significant. The custom data index (Model 1 and 2) and also the green part of the global aggregate index (Model 5 and 6) show a significant negative relationship. This might surprise first but considering that the *Payment Rank* is already included to some extent in the credit rating and that the majority of green bonds are senior unsecured the results could be driven by some outliers.

Overall, the results support our first hypothesis that having an *ESG rating* is rewarded with a lower spread, confirming a directional effect. Information asymmetry regarding

[5] In an alternative model, we have replaced the USD dummy variable with a variable controlling for bonds issued in Euro. Using this alternative variable does not change the results for the other variables. The results of this alternative model are not shown for reasons of brevity but are upon on request.

Table 6 Regression results Hypothesis 2

	Custom universe		Green bond index		Global aggregate index	
	Corporates	All	Corporates	All	Corporates	All
Model	1	2	3	4	5	6
Weighted average ESG score	− 13.570*** (3.426)	− 10.603*** (2.918)	− 6.091*** (1.811)	− 6.444*** (1.719)	− 9.634*** (1.701)	− 8.918*** (1.684)
Credit rating	14.128*** (3.280)	12.739*** (2.333)	4.669*** (1.006)	5.251*** (0.795)	7.886*** (1.020)	7.370*** (0.970)
Amount issued	− 17.469*** (6.064)	− 16.462*** (5.870)	− 16.951*** (3.589)	− 16.485*** (2.524)	− 16.698*** (4.043)	− 18.384*** (3.880)
Amount outstanding	− 2.745* (1.461)	− 1.921* (1.147)				
Time to maturity	25.535*** (4.096)	19.763*** (4.583)	32.125*** (3.055)	28.731*** (2.725)	27.967*** (3.927)	26.675*** (3.407)
Callable	− 11.059 (6.801)	− 3.450 (7.449)	− 0.905 (3.934)	0.384 (3.880)	− 9.797* (5.213)	− 8.559* (5.011)
USD	46.769*** (8.736)	52.413*** (8.469)	37.905*** (4.336)	41.350*** (4.106)	33.406*** (4.745)	34.858*** (4.447)
Governm. related		0.957 (7.896)		− 19.288*** (5.098)		− 9.467 (6.010)
Payment rank	0.996 (11.332)	7.830 (9.484)	39.400*** (4.027)	38.987*** (3.855)	− 1.326 (3.986)	− 0.625 (4.076)
Intercept	417.476*** (125.317)	353.842*** (113.736)	229.635*** (71.932)	223.608*** (49.807)	348.581*** (85.167)	380.701*** (81.310)
N	163	211	192	228	193	208
Adjusted R^2	0.65	0.64	0.76	0.78	0.74	0.74

Standard errors are reported in parentheses

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

the greenness of the green bond and the risk of financing greenwashing is reduced. The existence of an ESG rating leads to higher credibility of the company, represented through a more favourable spread when issuing green bonds. The results are in line with the majority of CSP/CEP—CFP research.

Does a higher ESG rating lead to a lower spread?

Following the result that having an *ESG rating* is correlated to a lower spread, we now only investigate green bonds that have an ESG rating to determine the magnitude effect and if the greenness of a green bond matters for its pricing. Table 6 shows the results for the model using the weighted average ESG score.

Looking at the ESG influence on the spread, for every single regression the *Weighted Average ESG Score* is statistically significant on the 1% level. An improvement in the ESG rating of 1 point (scale is 0–10) leads to a decrease of the spread by 6 to 13 bps. Correlation does not automatically imply causality, but the results indicate that the greenness of a green bond does matter. The greener the issuer, expressed by the ESG rating, the more an investor is ready to give up.

Unsurprisingly, the results of Table 6 show that the variable of *Credit Rating* has a statistically significant negative relationship to the bond spread. A one notch lower rating (e.g., BB + to BB) increases the spread by 5 to 14 bps depending on the sample. A higher *Amount Issued* is associated with a lower spread but the *Amount Outstanding* is only weakly statistically different from zero. The coefficient *Time to Maturity* is positive and strongly significant for all regressions with a 10% increase leading to a 2–3 bps higher spread. The dummy variable, *Callable*, has as for Hypothesis 1, a negative coefficient but lacks significance at the 5% level. Moreover, the relationship between the spread and the dummy variable *USD* has not changed. This result is significant for all regressions. As expected, the dummy variable *Government related* is negatively related to the spread but only in Model 4.

Looking at Hypothesis 2, we conclude that the higher the ESG rating, the lower the spread of green bonds. The results confirm our second hypothesis and provide evidence for the magnitude effect.

Table 7 Regression results Hypothesis 3

	Custom universe		Green bond index		Global aggregate index	
	Corporates	All	Corporates	All	Corporates	All
Model	1	2	3	4	5	6
E-score	−2.355 (1.926)	−2.703 (2.004)	−0.703 (0.966)	−0.541 (0.944)	−1.407 (1.034)	−1.209 (0.994)
S-score	−4.631* (2.514)	−2.451 (2.332)	−1.757 (1.169)	−2.117* (1.164)	−1.914 (1.313)	−2.066 (1.286)
G-score	−5.504*** (2.072)	−3.428** (1.567)	−2.730** (1.129)	−3.230*** (1.068)	−4.486*** (1.195)	−4.304*** (1.183)
Credit rating	13.445*** (3.373)	12.571*** (2.390)	4.240*** (0.943)	4.960*** (0.766)	7.286*** (1.063)	6.870*** (0.972)
Amount issued	−20.386*** (6.536)	−17.945*** (6.064)	−18.674*** (3.795)	−17.508*** (2.574)	−19.286*** (4.498)	−20.188*** (4.142)
Amount outstanding	−3.429* (1.741)	−1.816 (1.163)				
Time to maturity	25.025*** (4.234)	19.270*** (4.902)	31.753*** (3.069)	28.745*** (2.770)	26.575*** (4.044)	25.679*** (3.583)
Callable	−18.450** (7.436)	−7.836 (8.243)	−2.639 (4.013)	−2.028 (3.874)	−11.500** (5.477)	−10.751** (5.179)
USD	47.935*** (8.629)	53.908*** (8.464)	38.666*** (4.199)	42.202*** (4.011)	35.502*** (4.546)	36.915*** (4.293)
Governm. related		−4.223 (9.075)		−17.833*** (5.482)		−10.956* (5.906)
Payment rank	3.501 (10.590)	9.512 (9.209)	40.425*** (3.905)	39.957*** (3.714)	0.315 (4.021)	0.980 (4.064)
Intercept	485.699*** (132.926)	368.290*** (118.582)	259.018*** (76.161)	239.152*** (51.211)	389.780*** (94.151)	407.613*** (87.055)
N	163	211	192	228	193	208
Adjusted R^2	0.64	0.62	0.75	0.78	0.72	0.73

Standard errors are reported in parentheses

$^*p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$

For green bonds, environmental criteria should dominate social and governance criteria

We now separate the ESG rating into the E-, S- and G-score. Table 7 shows the regression results for the model including the three components.

We find that the E-Score coefficient is not significant in any model. The *S-Score* shows only weak significance in the first model but lacks significance in the other models, while the G-Score is strongly significant in all regressions.

This result is interesting, since it implies that the governance part of the ESG score is the main driver behind lower spreads, not the E-Score as expected. One possible explanation for this result could be the characteristic of a green bond itself. The main characteristic of a green bond is the use of proceeds. These need to be directed towards an environmentally friendly purpose. From an investors' point of view, the results indicate that the trustworthiness represented through the G-Score is more relevant than the environmental friendliness represented through the E-Score. Trust is a crucial point for green bonds due to the special use of proceeds. Therefore, the results suggest that the belief the issuer uses the proceeds in the stated way is more important than the environmental image of the issuer. We conclude that companies with high governance scores are able to issue green bonds more successfully.

We now return to the global aggregate index and our first question of the existence of a (negative) green bond premium. Running a correlation matrix,[6] it is obvious that the green bond characteristic shows almost no correlation with the other variables. This is particularly interesting for the E-Score. It indicates that bonds from environmental-friendly companies are not more likely to be green than bonds from less environmental-friendly firms. This supports the conclusion that the environmental friendliness of an issuer are not the most important characteristic for the spread of a green

[6] For reasons of brevity, the correlation matrix is not shown in the paper but is available upon request.

bond. We need to reject our third hypothesis, as a domination of E versus S and G could not be found. However, we find that governance seems to be the main driver of spreads of green bonds.

Conclusion

First of all, the evidence provided in our paper supports the (negative) green bond premium found in previous studies. For our sample, we report a negative premium of 8 to 14 basis points (directional effect). Addressing the question of potential greenwashing, we show that the existence of an ESG rating lowers the spread of green bonds. We offer the explanation of reduced information asymmetry and additional certification by the rating agency. Further, a higher ESG score (better rating) additionally lowers the spread for green bonds (magnitude effect). Remarkably, not the influence of the E-Score is the main driver for green bond spreads, but the G-Score (composition effect).

Our findings contribute in several ways to the literature: First, the ongoing existence of a negative green bond premium is underlined. Second, it is important for issuers to have an ESG rating in order to be investable for ESG-dedicated investors. Third, having a good rating does pay off spread-wise, as a good ESG rating will attract more investors. The shade of green matters for the pricing of green bonds. As the governance pillar has the strongest and most significant influence on the spread, we conclude that governance is an important driver of credit risk, even in the case of green bonds.

References

Bachelet, Maria J., Leonardo Becchetti, and Stefano Manfredonia. 2019. The green bonds premium puzzle: The role of issuer characteristics and third-party verification. *Sustainability* 11(4): 1098.

Barnett, Michael L., and Robert M. Salomon. 2012. Does it pay to be really good? Addressing the shape of the relationship between social and financial performance. *Strategic Management Journal* 33(11): 1304–1320.

Berg, Florian, Julian Kölbel, and Roberto Rigobon. 2019. Aggregate confusion: The divergence of ESG ratings. *SSRN Working Paper*. https://ssrn.com/abstract=3438533.

Capelle-Blancard, Gunther, and Stéphanie Monjon. 2012. Trends in the literature on socially responsible investment: Looking for the keys under the lamppost. *Business Ethics: A European Review* 21(3): 239–250.

Chatterji, Aaron K., David I. Levine, and Michael W. Toffel. 2009. How well do social ratings actually measure corporate social responsibility? *Journal of Economics & Management Strategy* 18(1): 125–169.

Chatterji, Aaron K., Rudolphe Durand, David I. Levine, and Samuel Touboul. 2016. Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Management Journal* 37(8): 1597–1614.

China Green Bond Market. 2019. Research Report. Climate Bonds Initiative and China Central Depository & Clearing Research Centre (CCDC Research). https://www.climatebonds.net/files/reports/china-sotm_cbi_ccdc_final_en260219.pdf.

Dorfleitner, Gregor, Gerhard Halbritter, and Mai Nguyen. 2015. Measuring the level and risk of corporate responsibility—An empirical comparison of different ESG rating approaches. *Journal of Asset Management* 16(7): 450–466.

Gatti, Stefano, and Andrea Florio. 2018. Issue spread determinants in the green bond market: The role of second party reviews and of the Green bond Principles. In *Research Handbook of Finance and Sustainability*. Edward Elgar Publishing.

Ge, Wenxia, and Mingzhi Liu. 2015. Corporate social responsibility and the cost of corporate bonds. *Journal of Accounting and Public Policy* 34(6): 597–624.

Gianfrate, Gianfranco, and Mattia Peri. 2019. The green advantage: Exploring the convenience of issuing green bonds. *Journal of Cleaner Production* 219: 127–135.

Hachenberg, Britta, and Dirk Schiereck. 2018. Are green bonds priced differently from conventional bonds? *Journal of Asset Management* 19(6): 371–383.

Kapraun, Julia, and Christopher Scheins. 2019. (In)-Credibly green: Which bonds trade at a green bond premium? *SSRN Working Paper*. https://doi.org/10.2139/ssrn.3347337.

Karpf, Andreas, and Antoine Mandel. 2017. Does it pay to be green? *SSRN Working Paper*. https://doi.org/10.2139/ssrn.2923484.

Kuhn, Dennis, Florian Kiesel, and Dirk Schiereck. 2018. Determinanten von Credit Spreads in Green bonds in europäischen Emissionsmärkten. *Zeitschrift für Umweltpolitik & Umweltrecht* 4: 422–436.

Laufer, William S. 2003. Social accountability and corporate greenwashing. *Journal of Business Ethics* 43(3): 253–261.

Maul, Daniel, and Dirk Schiereck. 2017. The bond event study methodology since 1974. *Review of Quantitative Finance and Accounting* 48(3): 749–787.

Menz, Klaus-Michael. 2010. Corporate social responsibility: Is it rewarded by the corporate bond market? A critical note. *Journal of Business Ethics* 96(1): 117–134.

Nollet, Joscha, George Filis, and Evangelos Mitrokostas. 2016. Corporate social responsibility and financial performance: A non-linear and disaggregated approach. *Economic Modelling* 52: 400–407.

Partridge, Candace and Francesca Medda. 2018. Green premium in the primary and secondary US municipal bond markets. *SSRN Working Paper*. https://doi.org/10.2139/ssrn.3237032.

Pierce, Jason R., and Herman Aguinis. 2013. The too-much-of-a-good-thing effect in management. *Journal of Management* 39(2): 313–338.

Polbennikov, Simon, Albert Desclée, Lev Dynkin, and Anando Maitra. 2016. ESG ratings and performance of corporate bonds. *The Journal of Fixed Income* 26(1): 21–41.

Preclaw, Ryan, and Anthony Bakshi. 2015. The cost of being green. *Barclays Research*. https://www.environmental-finance.com/assets/files/US_Credit_Focus_The_Cost_of_Being_Green.pdf, Accessed 20 Feb 2020.

Schiereck, Dirk, Gunnar Friede, and Alexander Bassen. 2019. Financial performances of green securities. In *The rise of green finance in Europe*, ed. M. Migliorelli and P. Dessertine. Cham: Palgrave Studies in Impact Finance, Palgrave Macmillan.

Semenova, Natalia, and Lars G. Hassel. 2015. On the validity of environmental performance metrics. *Journal of Business Ethics* 132(2): 249–258.

Stellner, Christoph, Christian Klein, and Bernhard Zwergel. 2015. Corporate social responsibility and Eurozone corporate bonds: The moderating role of country sustainability. *Journal of Banking & Finance* 59: 538–549.

Trumpp, Christoph, and Thomas Guenther. 2017. Too little or too much? Exploring U-shaped relationships between corporate environmental performance and corporate financial performance. *Business Strategy and the Environment* 26(1): 49–68.

Wu, Yue, Kaifu Zhang, and Jinhong Xie. 2020. Bad greenwashing, good greenwashing: Corporate social responsibility and information transparency. *Management Science* 66(7): 2801–3294.

Zerbib, Oliver D. 2019. The effect of pro-environmental preferences on bond prices: Evidence from green bonds. *Journal of Banking & Finance* 98: 39–60.

Moritz Immel He recently obtained his master's degree in business engineering (mechanical engineering with business administration) from Technical University of Darmstadt. He obtained his bachelor's degree also from TU Darmstadt in the same field of study. During his studies, he held several internships and working student roles in the field of banking and asset management in Frankfurt and London. Lastly, he was working student at Allianz Global Investors in Frankfurt. After he obtained his master's degree in April 2020, he joined Lupus alpha Asset Management AG in July 2020.

Britta Hachenberg She received her German diploma from the University of Applied Sciences, Cologne, Germany, in 2000. In the same year, she joined the investment bank Goldman Sachs and held various roles in Frankfurt and London as an Executive Director until she left the bank in 2016. In 2018, she finished her doctoral thesis at the Technical University of Darmstadt and started as a Professor at TH Koeln, Cologne. Hachenberg is a member of the advisory board of Koelner Boersenverein e.V. and focuses her research on capital markets and banking regulation.

Florian Kiesel He is Associate Professor of Finance at Grenoble Ecole de Management. He obtained his doctoral degree in Finance at the Technical University of Darmstadt, Germany. He previously studied Mechanical Engineering with Business Administration and worked for two years as Postdoc Researcher at the Chair of Corporate Finance, both at Technical University of Darmstadt. In 2016, he was visiting the Cass Business School in London, and, during his studies, he worked for international companies in Germany and China.

Dirk Schiereck He obtained his Master's Degree (German diploma) in Economics from Kiel University and his Ph.D. in Business Administration and Banking from Mannheim University, where he also finalized his habilitation in 2000. After 2 years at the Institute for Mergers and Acquisitions of Witten/Herdecke University, he spent 6 years at the European Business School. In 2008, he received an appointment for the Chair of Corporate Finance at Technical University of Darmstadt. Schiereck is a member of the Scientific Advisory Boards of DDV—Deutscher Derivate Verband (the German Derivatives Association) and DKS—Deutsche Kreditmarkt-Standards.

Journal of Asset Management (2021) 22:110–119
https://doi.org/10.1057/s41260-021-00206-4

ORIGINAL ARTICLE

Air pollution, investor sentiment and excessive returns

Matthew Muntifering[1]

Revised: 4 September 2020 / Accepted: 19 January 2021 / Published online: 1 March 2021

Abstract

This paper extends the asset pricing literature by offering a proprietary index of negative investor sentiment linked to carbon monoxide (CO), nitrogen dioxide (NO_2), ozone particle (O_3), 2.5 mm particulate matter ($PM_{2.5}$), and sulfur dioxide (SO_2) levels; determining the link between New York City air pollution and stock market returns. Kindly note that Food products and wholesale portfolio returns on average increase with enhanced negative investor sentiment. This is consistent with behaviors associated with psychological stress, like binge eating and shopping sprees. Personal services portfolio returns decrease on average with increased negative investor sentiment, consistent with behavioral isolationism.

Keywords asset pricing models · investor sentiment · air pollution

JEL Classification G11 · G12 · Q52

Introduction

How does air pollution affect the stock market? The objective of this paper is to assess the relationship between pollution, investor sentiment, and stock market returns. The United States stock market is the holding choice of over $30 trillion in wealth. A risk averse investor responds to uncertainty by his or her willingness to pay for a risk premium to achieve a more certain state. This is the fundamental underlying of the Capital Asset Pricing Model (CAPM) of Sharpe (1963). Their model says that the expected return of a risky asset can be explained by a composition of the difference between returns and risk-free rate of return plus the risk premium. Fama and French (1993) extend this relationship to include the difference between returns of portfolios diversified with small stocks and big stocks, respectively, and the difference between returns on high book-to-market value stocks versus low. Fama and French (2015) continue the extension by adding variables for the difference in return between highly profitable and the least profitable as well as one for the difference in returns for firms who invest aggressively versus conservatively. Investor sentiment may alter asset prices away from their fundamental value if only considering the predefined characteristics. This paper contributes a unique set of negative investor sentiment components to the asset pricing literature, generated using a principal-component analysis of New York City's daily carbon monoxide (CO), nitrogen dioxide (NO_2), ozone particle (O_3), 2.5 mm particulate matter ($PM_{2.5}$), and sulfur dioxide (SO_2) levels. Using pollution measures instead of weather variations are desirable because it is more feasible to implement a policy of emissions reduction than to control Mother Nature, and thus this paper provides an interesting alternative channel to explore in the welfare implications of emission reduction policies. The new model adds this index to the Fama and French (2015) five factors and is implemented on the 49 industry portfolios provided by Dr. French's website in order to examine if it can adequately explain returns across a wide variety of sectors and securities. The results indicate that sentiment based on pollution can further explain stock market returns and may be useful to implement into a trading strategy. The rest of the paper is organized as follows: a review of the relevant literature relating to stock market return factors, a theoretical derivation of the original CAPM of Sharpe (1963) and extension that serves the purpose of the paper, an overview of the data and empirical methodology, results, discussion, and conclusion.

Matthew Muntifering
mkm0042@auburn.edu

[1] Department of Agricultural Economics and Rural Sociology, Auburn University, 202 Comer Hall, Auburn, AL 36849, USA

Review of literature

Li et al. (2019) show a positive impact of air pollution on the disposition effect, which is a behavioral anomaly where traders hold onto assets whose prices are dropping and sell those who are increasing. They show that the effect is larger when measuring air pollution by 10 or 2.5 mm inhalable particulate matter (PM_{10} or $PM_{2.5}$), further validating the use of $PM_{2.5}$ data in this study. Levy and Yagil (2011), Lepori (2016), Li and Peng (2016), and An et al. (2018) also study the effects of air pollution on the stock market and find negative relationship between air pollution and returns in the United States, Italy, and China. Heyes et al. (2016) examine primarily $PM_{2.5}$ and finds a robust negative statistically significant relationship. The China AQI in Li and Peng (2016) contains information for carbon monoxide, nitrogen dioxide, ozone particle, 2.5 mm particulate matter, and sulfur dioxide. Their study highlights these as five associated with negative human health consequences, particularly $PM_{2.5}$ that can infiltrate alveoli and obstruct gas exchange.

Saunders (1993) examines the relationship between sunny days, investor sentiment and expected returns between 1983 and 1989 and finds no significant "sunshine effect." Hirshleifer and Shumway (2003) found that trading decisions made incorporating sunshine information can increase a portfolio's Sharpe ratio, measuring excess returns for a given unit of risk (measured in standard deviation of returns). The Britten-Jones test can be used to identify this relationship and is conducted by regression of 1's on the vector of portfolio returns. They do mention however that these results are sensitive to the frequency of trades if non-trivial transaction costs exist. Their study examines daily market returns in 26 countries from 1982 to 1997. Their use of sunshine in the city where the most active financial exchange exists motivates the use of using New York City weather data. Chang et al. (2008) conclude that increased cloud cover in New York City is associated with increased stock volatility. Their results confirm those of Saunders (1993) and Hirshleifer and Shumway (2003) in that there are linkages between daily weather patterns and asset returns. Trombley (1997) critiques Saunders (1993) paper by saying the distribution of cloudiness lends itself to statistically significant results by comparing 20% cloudiness to 100%. We address this concern by using widely dispersed pollution data. Cao and Wei (2005) examine temperature effects on stock market returns while remaining consistent with the trend in the literature to include data from the major market city. Kliger and Levy (2003) find that increased cloud cover is related to increased investor perceived probabilities of negative events. They find that higher temperatures are associated with apathy and lower returns while adjacently lower temperatures imply aggression and more risk seeking and higher returns. Dowling and Lucey (2008a, b) study wind, precipitation, and geomagnetic storms and find that they are relevant drivers of increased volatility of individual indices. Loughran and Schultz (2004) find that blizzards and cloudy days in New York are associated with marginally lower stock returns. One interesting question arises when considering cloud covers studies. How much of the cloud cover can be attributed to air pollution? While the literature appears deeper regarding the effects of weather rather than pollution on the stock market, the detrimental physical health effects of pollution in addition to the adverse mental health effects from clouds that could also be from pollution signal an importance for the field to continue to understand these relationships with the goal to improve human welfare across the multiple aforementioned avenues.

Jaffe et al. (1989), Wang et al. (1997) and Pettengill (2003) examine what is referred to as the Monday effect, which originally said there is a decline in labor productivity on this day relative to others. It has been shown to have the reversal effect. Schultz (1985), Ariel (1987), and Kamstra et al. (2000a) describe another seasonal effect pervasive in the literature, the January effect. This effect says that unusually high returns are observed in January relative to other months.

Other examples of seasonal occurrences that may explain returns are Kamstra et al. (2000b) and Kamstra et al. (2003) who find that daylight savings time and thus shortened days lead to seasonal affective disorder (SAD), associated with depression, which causes an increase in risk aversion. This heightened risk aversion thus leads to decreased variability or volatility in asset returns. Dowling and Lucey (2008a, b) also examine daylights saving time and lunar phases with similar results. Loughran and Schultz (2004) in addition to their weather results find that trading is slowed in cities with high Jewish populations on Yom Kippur. Also, Dichev and Janes (2003), Yuan et al. (2006) and Keef and Khaled (2011) among others study lunar phases of the moon and stock market returns and provide evidence for moon effects. These studies support the findings that returns are higher on new moon days but argue that it is difficult to imagine that, the Monday effect, or the turn-of-the-month effect as significant drivers off inefficient markets. This paper agrees and argues that as more recent technological advances such as algorithmic trading strategies, near-zero cost investment platforms and free financial literacy training mobile applications continue to progress, the increased financial savvy of investors across investors of all skill levels will increasingly diminish these particular inefficiencies and thus do not consider these variables.

Theoretical framework

The theoretical model relies on the fundamental equation of the Sharpe (1963) CAPM:

$$E(r_j) = a_j + r + [E(r_m) - r]\beta_j, \quad \text{for each } j. \tag{1}$$

The standard assumption is that if the model adequately explains returns then alpha should be zero. The arguments against the efficient market hypothesis can be found in Barberis et al. (2005). They find that asset returns are indeed influenced by non-common fundamental risk. These deviations from the assumption provide the Fama and French (1992) framework and allow for further analysis.

The Three Factor Model of Fama and French (1993) includes size (SMB) and book-to-market value (HML) factors to the original CAPM specification. They find significant explanatory power of these factors towards excess return, stating the book-to-market value is positively correlated with asset returns. The same holds true for firm size as measured by market capitalization.

$$E(r_j) - r = a_j + [E(r_m) - r]\beta_{j1} + \mathrm{SMB}\beta_{j2} + \mathrm{HML}\beta_{j3}, \quad \text{for each } j. \tag{2}$$

Fama and French (2016) extend the model further to the Five Factor Model to include profitability (RMW) and an investment aggressiveness (CMA) factor. They also note that when implementing this model, the value (book-to-market/HML) factor may become redundant. As an additional avenue of analysis, this paper explores the issue further.

$$\begin{aligned} E(r_j) - r = {} & a_j + [E(r_m) - r]\beta_{j1} + \mathrm{SMB}\beta_{j2} \\ & + \mathrm{HML}\beta_{j3} + \mathrm{RMW}\beta_{j4} + \mathrm{CMA}\beta_{j5}, \quad \text{for each } j. \end{aligned} \tag{3}$$

By asserting that investor sentiment may further explain excess asset returns, this paper extends the theoretical model to include a sixth factor, a unique measure of investor sentiment (SEN).

$$\begin{aligned} E(r_j) - r = {} & a_j + [E(r_m) - r]\beta_{j1} + \mathrm{SMB}\beta_{j2} + \mathrm{HML}\beta_{j3} \\ & + \mathrm{RMW}\beta_{j4} + \mathrm{CMA}\beta_{j5} + \mathrm{SEN}\beta_{j6}, \quad \text{for each } j. \end{aligned} \tag{4}$$

The parameters of the model can be estimated using a regression like Fama and MacBeth (1973).

Data

Daily air quality index (AQI) data for carbon monoxide, nitrogen dioxide, ozone particle, 2.5 mm particulate matter, and sulfur dioxide in New York City were obtained from the Environmental Protection Agency on trading days from January 4, 2013 until May 7, 2019. Data across a similar time horizon on portfolio returns and the Fama and French (2016) factors are obtained from Dr. French's personal website. The number of trading days in this sample is 1,617. The period before 2013 is excluded from the analysis in order to avoid overlap into the financial crisis.

Table 1 in the "Appendix" describes the summary statistics of the dataset. The portfolios were constructed based on the stock's industry SIC code. The portfolio types are; agriculture, food products, candy and soda, beer and liquor, tobacco products, recreation, entertainment, printing and publishing, consumer goods, apparel, healthcare, medical equipment, pharmaceutical drugs, chemicals, rubber and plastic products, textiles, construction materials, construction, steel works, fabricated products, machinery, electrical equipment, automobiles and trucks, aircraft, shipbuilding and railroad equipment, defense, precious metals, non-metallic and industrial mining metal, coal, petroleum and natural gas, utilities, communication, personal services, business services, computer hardware, computer software, electronic equipment, measuring and control equipment, business supplies, shipping containers, transportation, wholesale, retail, restaurants and hotels/motels, banking, insurance, real estate, trading, and other. Figures 1, 2, 3, 4 and 5 show the normalized AQI values of various pollutants over time. Visually there is sparsity in the data and alleviates the concern pressed in Trombley (1997).

Empirical framework

Principal component analysis is used to create the sentiment components. A Bartlett test of sphericity with a null that the correlation matrix for the given variables is not an identity matrix is rejected. The Kaiser–Meyer–Olkin Measure of Sampling Adequacy is 0.643 which is greater than 0.5. The diagonal on the anti-image correlation coefficient matrix can be interpreted as a measuring of sampling adequacy. The values for this diagonal are 0.66, 0.6989, 0.8321, 0.6487, and 0.8113, respectively. Further, low values along the diagonal of the residual correlation matrix are a measure of performance of the components in explaining the variation in the original data. These values are all under 0.0005. These results taken in conjunction imply that a principal component analysis of the data may be appropriate.

The first two principal components have eigenvalues above 1 and represent the points to the left of the "bow" of the scree plot (Figure 6), which are two useful criteria in deciding the number of components to use. These components explain roughly 67 percent of the variation in the initial data and may be adequate for analysis (Table 2).

This technique is used in the asset pricing literature to develop indices of investor sentiment by Baker and Wurgler (2006), Lin et al. (2012), Ait-Sahalia and Xiu (2017), Dhaoui and Bensalah (2017) and Gerber et al. (2019). Liew and Budavari (2017) use a combination of proprietary StockTwits data to construct their index and provide another unique example of how to measure sentiment.

The loading plot (Figure 7) and loading table (Table 3) show the makeup of the individual components. Component 1 is as followed (recalling that the input data has been normalized):

$$\begin{aligned}\mathrm{COMP1} &= .5145\mathrm{CO} + .5207\mathrm{NO_2} + .0917\mathrm{OZONE} \\ &+ .5256\mathrm{PM_{2.5}} + .4237\mathrm{SO_2}\end{aligned} \tag{5}$$

Given the theoretical model and the derivation of principal components, the final model to be estimated is:

$$\begin{aligned}E(r_j) - r = a_j + [E(r_m) - r]\beta_{j1} + \mathrm{SMB}\beta_{j2} + \mathrm{HML}\beta_{j3} \\ + \mathrm{RMW}\beta_{j4} + CMA\beta_{j5} + \mathrm{COMP1}\beta_{j6}, \quad \text{for each} j.\end{aligned} \tag{6}$$

Results and discussion

The analysis was performed using Stata 15. The estimated output tables can be found in Table 4.

The results indicate that the investor sentiment index composition in this paper do indeed help explain stock market returns. The 49 industry portfolios examined all display an increase in the R-square value as a result of the implementation. Further, the F-test for the null hypothesis that the factors are jointly insignificant is rejected. The portfolios where the first principal component is individually statistically significant are food products (positive effect), personal services (negative effect), and wholesale (positive). These portfolios correspond to actions in psychological and health literature taken in companionship with stress. The psychological link to binge eating, isolationism, and shopping sprees are well documented. see Smith et al. (1998), Sanders et al. (2000) and Krueger (1988), respectively. It is also of interest that statistically significant positive alpha was generated for guns, business services, and insurance portfolios.

Conclusion

These results add to the expanding literature on the effects of pollution on stock market returns, specifically providing a link between New York City pollution levels and excessive returns. A unique measure of negative investor sentiment, generated by using a principal-component analysis of New York City's daily carbon monoxide, nitrogen dioxide, ozone, 2.5 mm particulate matter, and sulfur dioxide levels is also contributed. Further research into this subfield may include adding similar pollution variables or investor sentiment data to the index. Finally, this paper gives supporting evidence to emission reduction policies, giving an alternative vehicle to welfare improvement because of successful implementation.

Appendix 1: Tables

See Tables 1, 2, 3 and 4.

Table 1 Descriptive statistics

Variable	Obs	Mean	Std.dev.	Min	Max
coaqi	1617	8.772	3.603	2	38
no2aqi	1617	42.244	13.714	15	131
ozoneaqi	1617	47.952	25.719	11	210
pm25aqi	1617	50.838	15.93	16	141
so2aqi	1617	4.989	5.056	0	59
mrf	1617	0.051	0.832	−4.03	5.06
smb	1617	−0.007	0.493	−1.63	2.52
hml	1617	−0.014	0.494	−1.69	2.38
rmw	1617	0.004	0.324	−1.58	1.63
cma	1617	−0.009	0.316	−1.32	1.96
rf	1617	0.002	0.003	0	.01
agric	1617	0.022	1.122	−6.43	7.65
food	1617	0.038	0.882	−5.11	4.73
soda	1617	0.043	0.882	−7.3	5.29
beer	1617	0.058	0.825	−4.28	3.04
smoke	1617	0.038	1.056	−11.46	4.91
toys	1617	0.04	1.391	−8.15	8.22
fun	1617	0.088	1.464	−6.75	7.15
books	1617	0.024	1.128	−8.72	6.91
hshld	1617	0.043	0.806	−3.98	4.55
clths	1617	0.057	1.156	−6.3	6.36
hlth	1617	0.045	1.147	−8.86	4.78
medeq	1617	0.076	0.972	−4.45	4.92
drugs	1617	0.054	1.044	−4.68	6.25
chems	1617	0.04	1.061	−4.75	5.37
rubbr	1617	0.053	0.987	−5.11	3.81
txtls	1617	0.048	1.409	−18.31	6.56
bldmt	1617	0.044	1.129	−5.21	4.17
cnstr	1617	0.039	1.307	−6.2	5.13
steel	1617	0.019	1.586	−6.85	8.72
fabpr	1617	0.035	1.743	−15.45	9.78
mach	1617	0.045	1.143	−5.93	5.23
elceq	1617	0.033	1.104	−4.82	4.95
autos	1617	0.032	1.219	−5.98	5.29
aero	1617	0.075	1.065	−5.36	5.12
ships	1617	0.063	1.392	−5.64	7.99
guns	1617	0.097	1.042	−5.66	6.22
gold	1617	0.003	2.405	−11.76	10.42
mines	1617	.006	1.642	−7.52	10.04
coal	1617	−0.055	2.799	−18.44	18.08
oil	1617	0.006	1.326	−7.47	6.71
util	1617	0.044	0.835	−4.44	2.89
telcm	1617	0.045	0.851	−4.42	3.6
persv	1617	0.045	1.091	−4.47	4.35
bussv	1617	0.075	0.952	−4.4	5.17
hardw	1617	0.049	1.169	−7	5.11
softw	1617	0.074	1.092	−4.8	6.33
chips	1617	0.081	1.17	−7.25	5.91
labeq	1617	0.076	1.048	−4.56	5.02
paper	1617	0.044	0.925	−7.35	4.15

Table 1 (continued)

Variable	Obs	Mean	Std.dev.	Min	Max
boxes	1617	0.05	1.09	−5.55	4.4
trans	1617	0.059	1.083	−4.66	5.61
whlsl	1617	0.041	0.894	−4.18	3.85
rtail	1617	0.063	0.93	−4.1	6.88
meals	1617	0.065	0.837	−4.31	3.84
banks	1617	0.059	1.096	−6.17	5.07
insur	1617	0.071	0.88	−4.37	4.26
rlest	1617	0.025	1.16	−7.25	6.13
fin	1617	0.06	1.158	−6.92	5.23
other	1617	0.036	0.887	−5.25	4.65
zco	1617	0	1	−1.879	8.112
zno2	1617	0	1	−1.987	6.472
zozone	1617	0	1	−1.437	6.301
zpm25	1617	0	1	−2.187	5.66
zso2	1617	0	1	−.987	10.683
pc1	1617	0	1.466	−3.377	8.737
pc2	1617	0	1.083	−4.123	6.281
pc3	1617	0	0.841	−4.067	7.986
pc4	1617	0	0.74	−3.124	5.331
pc5	1617	0	0.65	−2.022	5.19

Table 2 Proportion of variance explained by components

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	2.148	0.975	0.430	0.430
Comp2	1.174	0.466	0.235	0.664
Comp3	0.708	0.161	0.142	0.806
Comp4	0.547	0.124	0.110	0.915
Comp5	0.423		0.085	1.000

Table 3 Loading table

Variable	Comp1	Comp2	Comp3	Comp4	Comp5	Unex-plained
zco	0.514	−0.264	−0.436	−0.337	0.602	0
zno2	0.521	0.098	−0.362	0.721	−0.261	0
zozone	0.092	0.858	0.176	0.075	0.467	0
zpm25	0.526	0.277	0.112	−0.564	−0.563	0
zso2	0.424	−0.328	0.797	0.206	0.186	0

Table 4 Regression results

	agric	food	soda	beer	smoke	toys
mrf	0.771***	0.823***	0.686***	0.724***	0.733***	0.990***
smb	0.0684	−0.177***	−0.353***	−0.368***	−0.329***	0.433***
hml	0.0439	−0.266***	−0.284***	−0.303***	−0.319***	−0.217**
rmw	0.137	0.416***	0.423***	0.422***	0.558***	0.331***
cma	0.178	0.803***	0.738***	0.633***	0.804***	0.252*
pc1	0.0280	0.0302**	−0.00342	0.00232	0.00774	−0.00390
_cons	−0.0155	−0.00306	0.00659	0.0187	−0.00119	−0.00947
N	1617	1617	1617	1617	1617	1617
R^2	0.308	0.506	0.373	0.476	0.296	0.375
	fun	books	hshld	hlth	medeq	drugs
mrf	1.193***	1.027***	0.809***	0.849***	0.895***	0.936***
smb	0.156**	0.637***	−0.230***	0.373***	0.0342	−0.0859**
hml	−0.339***	0.0959*	−0.276***	−0.263***	−0.589***	−0.641***
rmw	−0.284***	0.284***	0.409***	0.0336	−0.182***	−0.570***
cma	−0.348***	0.225**	0.822***	0.139	0.187***	0.314***
pc1	−0.0218	−0.0101	0.00382	0.00608	0.00123	0.00391
_cons	0.0211	−0.0218	0.00214	0.00186	0.0250	0.00175
N	1617	1617	1617	1617	1617	1617
R^2	0.584	0.659	0.587	0.448	0.713	0.716
	chems	rubbr	txtls	bldmt	cnstr	steel
mrf	1.106***	0.965***	1.094***	1.170***	1.180***	1.379***
smb	0.145***	0.389***	0.363***	0.587***	0.647***	0.758***
hml	0.172***	−0.0959**	−0.103	0.118***	0.232***	0.552***
rmw	0.122**	0.236***	0.391***	0.292***	0.246***	0.0429
cma	0.287***	0.278***	−0.00806	0.354***	0.218**	0.581***
pc1	0.00651	0.00375	0.0222	−0.00296	0.0114	−0.00451
_cons	−0.0110	0.00683	−0.00813	−0.00777	−0.0122	−0.0327
N	1617	1617	1617	1617	1617	1617
R^2	0.725	0.689	0.439	0.801	0.636	0.622
	fabpr	mach	elceq	autos	aero	ships
mrf	1.234***	1.222***	1.162***	1.159***	1.077***	1.198***
smb	1.088***	0.339***	0.395***	0.415***	−0.0168	0.686***
hml	0.472***	0.252***	0.133***	0.271***	−0.0112	0.232***
rmw	0.121	0.206***	0.151***	0.229***	0.308***	0.197**
cma	0.181	0.251***	0.423***	−0.0223	0.302***	0.435***
pc1	0.0145	−0.00198	0.00320	−0.00950	0.00287	0.0212
_cons	−0.0122	−0.00997	−0.0180	−0.0216	0.0208	0.0136
N	1617	1617	1617	1617	1617	1617
R^2	0.488	0.798	0.781	0.667	0.637	0.580
	guns	gold	mines	coal	oil	util
mrf	0.836***	0.448***	1.332***	1.368***	1.151***	0.616***
smb	−0.0451	0.193	0.401***	0.982***	−0.0359	−0.243***
hml	−0.262***	−0.183	0.519***	1.054***	0.514***	−0.220***
rmw	0.323***	−0.314	−0.0821	−0.629**	−0.755***	0.214***
cma	0.451***	1.207***	0.551***	0.852***	0.754***	0.733***
pc1	−0.00557	−0.0239	−0.0120	−0.0133	−0.00600	0.00955
_cons	0.0531**	−0.00829	−0.0465	−0.0926	−0.0358	0.0134

Table 4 (continued)

	guns	gold	mines	coal	oil	util
N	1617	1617	1617	1617	1617	1617
R^2	0.395	0.038	0.497	0.272	0.629	0.324
	telcm	persv	bussv	hardw	softw	chips
mrf	0.878***	1.039***	1.050***	1.104***	1.047***	1.093***
smb	−0.0551*	0.577***	0.186***	0.0597	−0.126***	0.00641
hml	−0.0347	0.0882*	−0.170***	−0.0661	−0.288***	−0.0998*
rmw	0.272***	0.240***	0.0752**	0.0737	−0.137***	0.310***
cma	0.441***	0.191**	−0.0158	−0.189*	−0.729***	−0.710***
pc1	0.00214	−0.0208*	−0.00124	−0.0147	−0.00748	−0.000746
_cons	0.00196	−0.00179	0.0201**	−0.0103	0.00918	0.0161
N	1617	1617	1617	1617	1617	1617
R^2	0.639	0.708	0.899	0.651	0.858	0.713
	labeq	paper	boxes	trans	whlsl	rtail
mrf	1.121***	1.029***	1.037***	1.133***	0.964***	0.982***
smb	0.0932***	−0.00293	0.203***	0.233***	0.326***	0.00937
hml	−0.297***	−0.0572	0.0636	0.117**	−0.0271	−0.289***
rmw	−0.0845*	0.419***	0.279***	0.436***	0.192***	0.452***
cma	0.171***	0.504***	0.425***	0.196**	0.327***	0.0247
pc1	−0.00295	−0.000334	−0.00471	0.0154	0.0130*	0.00675
_cons	0.0177	−0.00635	0.00183	0.00393	−0.00459	0.00710
N	1617	1617	1617	1617	1617	1617
R^2	0.835	0.744	0.587	0.718	0.811	0.765
	meals	banks	insur	rlest	fin	other
mrf	0.821***	1.059***	0.922***	1.098***	1.178***	0.959***
smb	−0.0271	0.0913***	−0.0377	0.446***	0.0909***	−0.203***
hml	−0.159***	1.025***	0.339***	0.125**	0.765***	0.274***
rmw	0.259***	−0.327***	−0.0614	0.193***	−0.329***	−0.108**
cma	0.124*	−0.718***	−0.103*	0.0348	−0.403***	0.255***
pc1	−0.00404	−0.0117	−0.00114	0.0108	−0.0139	0.00631
_cons	0.0210	0.0142	0.0274**	−0.0270	0.00898	−0.00795
N	1617	1617	1617	1617	1617	1617
R^2	0.625	0.879	0.780	0.674	0.859	0.779

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Appendix 2: Figures

See Figures 1, 2, 3, 4, 5, 6 and 7.

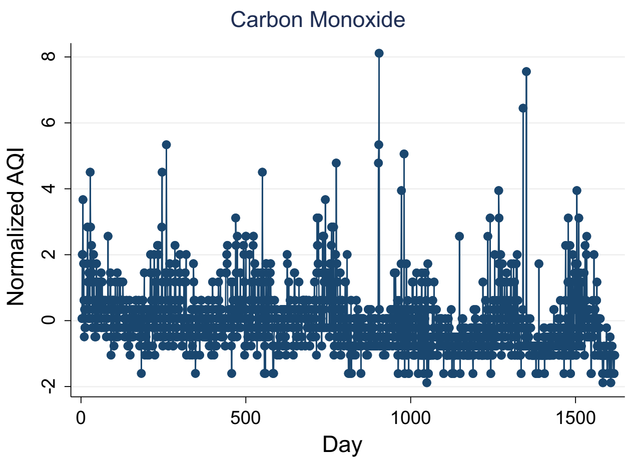

Fig. 1 Carbon monoxide

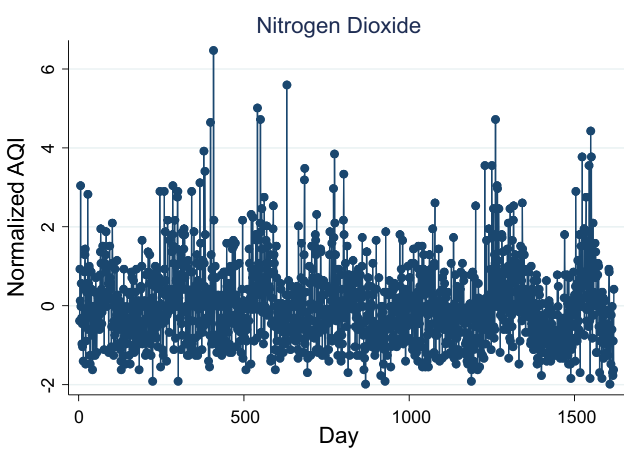

Fig. 2 Nitrogen dioxide

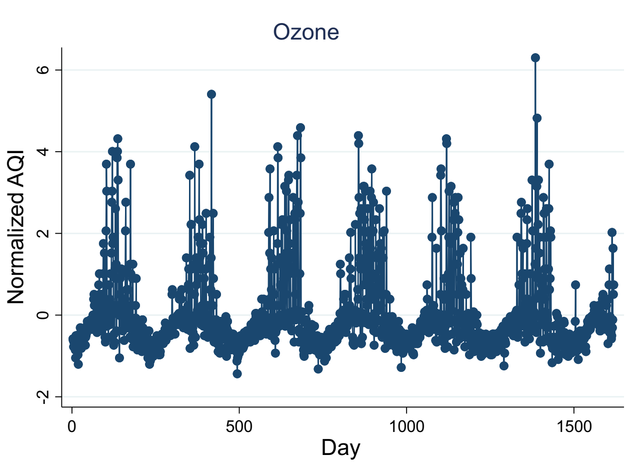

Fig. 3 Ozone

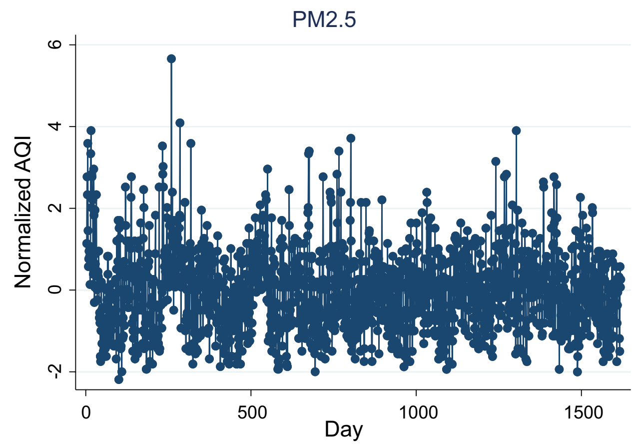

Fig. 4 $PM_{2.5}$

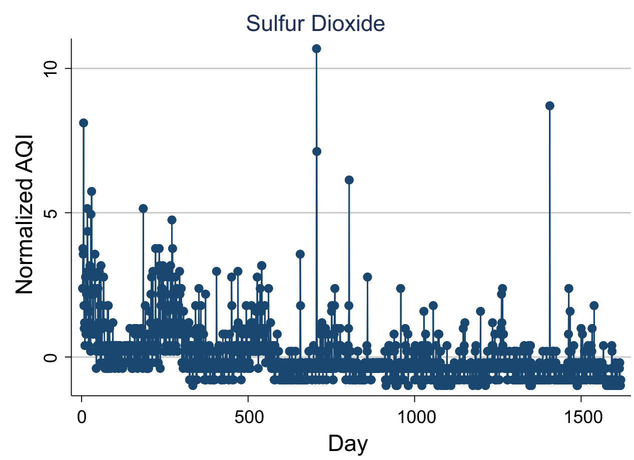

Fig. 5 Sulfur dioxide

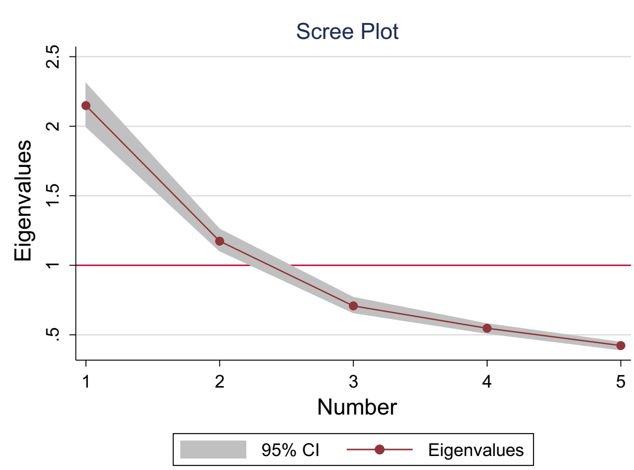

Fig. 6 Scree plot

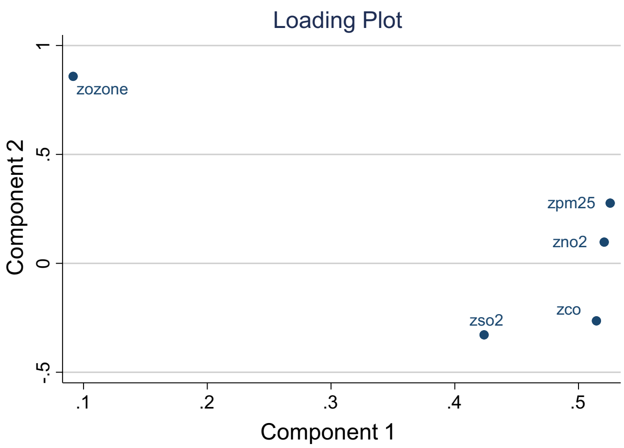

Fig. 7 Loading plot

References

Ait-Sahalia, Yacine, and Dacheng Xiu. 2017. Using principal component analysis to estimate a high dimensional factor model with high-frequency data. *Journal of Econometrics* 201 (2): 384–399.

An, N., B. Wang, P. Pan, K. Guo, and Y. Sun. 2018. Study on the influence mechanism of air quality on stock market yield and volatility: Empirical test from China based on GARCH model. *Finance Research Letters* 26: 119–125.

Ariel, Robert A. 1987. A monthly effect in stock returns. *Journal of Financial Economics* 18 (1): 161–174.

Baker, Malcolm, and Jeffrey Wurgler. 2006. Investor sentiment and the cross-section of stock returns. *The Journal of Finance* 61 (4): 1645–1680.

Barberis, Nicholas, Andrei Shleifer, and Robert W. Vishny. 2005. A model of investor sentiment. *Advances in behavioral finance* 2: 423–459.

Cao, M., and J. Wei. 2005. Stock market returns: A note on temperature anomaly. *Journal of Banking & Finance* 29 (6): 1559–1573.

Chang, S.C., S.S. Chen, R.K. Chou, and Y.H. Lin. 2008. Weather and intraday patterns in stock returns and trading activity. *Journal of Banking & Finance* 32 (9): 1754–1766.

Dhaoui, Abderrazak, and Nesrine Bensalah. 2017. Asset valuation impact of investor sentiment: A revised Fama-French five-factor model. *Journal of Asset Management* 18 (1): 16–28.

Dichev, Ilia D., and Troy D. Janes. 2003. Lunar cycle effects in stock returns. *The Journal of Private Equity* 6 (4): 8–29.

Dowling, Michael, and Brian M. Lucey. 2008a. Robust global mood influences in equity pricing. *Journal of Multinational Financial Management* 18 (2): 145–164.

Dowling, Michael, and Brian M. Lucey. 2008b. Mood and UK equity pricing. *Applied Financial Economics Letters* 4 (4): 233–240.

Fama, Eugene F., and Kenneth R. French. 1992. The cross-section of expected stock returns. *The Journal of Finance* 47 (2): 427–465.

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

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

Fama, Eugene F., and James D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81 (3): 607–636.

Gerber, Sander, Babak Javid, Harry Markowitz, Paul Sargen, and David Starer. 2019. A non-normal principal components model for security returns. *Available at SSRN 3352062*.

Heyes, Anthony, Matthew Neidell, and Soodeh Saberian. 2016. *The effect of air pollution on investor behavior: Evidence from the S&P 500*. No. w22753. National Bureau of Economic Research.

Hirshleifer, D., and T. Shumway. 2003. Good day sunshine: Stock returns and the weather. *The Journal of Finance* 58 (3): 1009–1032.

Jaffe, Jeffrey F., Randolph Westerfield, and Christopher Ma. 1989. A twist on the Monday effect in stock prices: Evidence from the US and foreign stock markets. *Journal of Banking & Finance* 13 (4–5): 641–650.

Kamstra, Mark, Lisa A. Kramer, and Maurice D. Levi. 2000a. Winter Blues: Seasonal Affective Disorder (SAD), the January Effect and the Stock Market. Simon Fraser University, Department of Economics.

Kamstra, Mark J., Lisa A. Kramer, and Maurice D. Levi. 2000b. Losing sleep at the market: The daylight saving anomaly. *American Economic Review* 90 (4): 1005–1011.

Kamstra, Mark J., Lisa A. Kramer, and Maurice D. Levi. 2003. Winter blues: A SAD stock market cycle. *American Economic Review* 93 (1): 324–343.

Keef, Stephen P., and Mohammed S. Khaled. 2011. Are investors moonstruck? Further international evidence on lunar phases and stock returns. *Journal of Empirical Finance* 18 (1): 56–63.

Kliger, Doron, and Ori Levy. 2003. Mood-induced variation in risk preferences. *Journal of Economic Behavior & Organization* 52 (4): 573–584.

Krueger, David W. 1988. On compulsive shopping and spending: A psychodynamic inquiry. *American Journal of Psychotherapy* 42 (4): 574–584.

Lepori, Gabriele M. 2016. Air pollution and stock returns: Evidence from a natural experiment. *Journal of Empirical Finance* 35: 25–42.

Levy, Tamir, and Joseph Yagil. 2011. Air pollution and stock returns in the US. *Journal of Economic Psychology* 32 (3): 374–383.

Li, Jennifer, Massimo Massa, Hong Zhang, and Jian Zhang. 2019. Air pollution, behavioral bias, and the disposition effect in China. *Journal of Financial Economics*. https://doi.org/10.1016/j.jfineco.2019.09.003.

Li, Q., and C.H. Peng. 2016. The stock market effect of air pollution: evidence from China. *Applied Economics* 48 (36): 3442–3461.

Liew, Jim, and Tamas Budavari. 2017. The “sixth” factor—A social media factor derived directly from tweet sentiments. *The Journal of Portfolio Management* 43 (3): 102–111.

Lin, Jianhao, Meijin Wang, and Lingfeng Cai. 2012. Are the Fama-French factors good proxies for latent risk factors? Evidence from the data of SHSE in China. *Economics Letters* 116 (2): 265–268.

Loughran, T., and P. Schultz. 2004. Weather, stock returns, and the impact of localized trading behavior. *Journal of Financial and Quantitative Analysis* 39 (2): 343–364.

Pettengill, Glenn N. 2003. A survey of the Monday effect literature. *Quarterly Journal of Business and Economics* 42 (3/4): 3–27.

Sanders, Christopher E., Tiffany M. Field, Miguel Diego, and Michele Kaplan. 2000. The relationship of Internet use to depression and social isolation among adolescents. *Adolescence* 35 (138): 237–237.

Saunders, E.M. 1993. Stock prices and Wall Street weather. *The American Economic Review* 83 (5): 1337–1345.

Schultz, Paul. 1985. Personal Income Taxes and the January Effect: Small Firm Stock Returns Before the War Revenue Act of 1917: A Note. *The Journal of Finance* 40 (1): 333–343.

Sharpe, William F. 1963. A Simplified Model for Portfolio Analysis. *Management Science* 9 (2): 277–293.

Smith, Delia E., Marsha D. Marcus, Cora E. Lewis, Marian Fitzgibbon, and Pamela Schreiner. 1998. Prevalence of binge eating disorder, obesity, and depression in a biracial cohort of young adults. *Annals of Behavioral Medicine* 20 (3): 227–232.
Trombley, Mark A. 1997. Stock prices and Wall Street weather: Additional evidence. *Quarterly Journal of Business and Economics* 36 (3): 11–21.
Wang, Ko., Yuming Li, and John Erickson. 1997. A new look at the Monday effect. *The Journal of Finance* 52 (5): 2171–2186.
Yuan, Kathy., Lu. Zheng, and Qiaoqiao Zhu. 2006. Are investors moonstruck? Lunar phases and stock returns. *Journal of Empirical Finance* 13 (1): 1–23.

Matthew Muntifering is a Ph.D. candidate and course instructor at Auburn University in the Department of Agricultural Economics and Rural Sociology.

Journal of Asset Management (2021) 22:120–132
https://doi.org/10.1057/s41260-020-00202-0

ORIGINAL ARTICLE

Sustainability efforts, index recognition, and stock performance

Moonsoo Kang[1] · K. G. Viswanathan[2] · Nancy A. White[2] · Edward J. Zychowicz[2]

Revised: 11 December 2020 / Accepted: 14 December 2020 / Published online: 12 January 2021

Abstract
We examine the long-term performance of stocks appearing in the Dow Jones Sustainability Index North America. We find that sustainability stocks exhibit abnormal returns for 12–30 months after the index listing, while those stocks generate no excess returns before the index listing. Moreover, sustainability stocks experience an increase in institutional ownership after the index listing. However, we find no evidence that short sellers increase their position to exploit a possible overpricing for sustainability stocks. Overall, our analysis suggests that sustainability efforts translate into a permanent increase in demand for stocks, leading to the superior performance.

Keywords Sustainable investing · ESG · Stock performance · Institutional ownership · Short sales

Introduction

According to the U.N. World Commission on Environment and Development, "sustainability is to meet the needs of the present without compromising the ability of future generations to meet their own needs." The Union of Conservation Scientists (IUCN), the United Nations Environment Programme (UNEP), and the World Wide Fund for Nature (WWF) view sustainability as "improving the quality of human life while living within the carrying capacity of the Earth's supporting eco-systems." Moreover, Krosinsky et al. (2011) suggest that sustainable investing (SI) goes beyond more established values-based SRI (socially responsible investing). SI, in their view, proactively takes into explicit consideration the impact of decisions on global issues such as pollution, resource depletion, and population growth. Such lofty aspirations have been embraced by individuals, governments, organizations, and businesses worldwide. Indeed, a McKinsey (2014) survey revealed that 36% of CEOs view sustainability as a top 3 priority.

However, the linkage between the popular acceptance of sustainability in practice and the evidence of actual corporate value creation is still underexplored in the literature (Kitzmueller and Shimshack 2012). Existing empirical evidence shows that the costs and benefits associated with sustainability efforts are mixed. On the one hand, sustainability efforts can enhance firm value by providing a form of insurance against adverse events and/or product market differentiation (Servaes and Tamayo 2013; Eccles et al. 2014; Hong and Liskovich 2016; Lins et al. 2017; Albuquerque et al. 2019). On the other hand, an increase in sustainability efforts can be an agency cost of firms. If firm's managers care about outside stakeholders or enjoy private benefits from sustainability investments, they can incur the cost of a sustainability commitment, regardless of operating performance (Masulis and Reza 2015; Cheng et al. 2016; Cronqvist and Yu 2017).[1] In this study, we contribute to ongoing efforts to fill this lacuna by examining the market reaction to firms that have been *first identified* as sustainability-consistent organizations. Specifically, we attempt to provide supporting evidence for the positive impact of sustainability efforts on a firm's market performance by examining whether firms with

We thank the S&P Dow Jones Indices and SAM for providing us with the Dow Jones Sustainability Index North America constituent data.

✉ Moonsoo Kang
moonsoo.kang@brooklyn.cuny.edu

1 Brooklyn College of the City University of New York, 2900 Bedford Ave, Brooklyn, NY 11210, USA

2 Hofstra University, Hempstead, NY, USA

[1] Cronqvist and Yu (2017) develop a simple theoretical framework involving a utility-maximizing CEO with social preferences to predict a CEO-daughter effect in the context of corporate decision-making with respect to stakeholders other than a firm's shareholders. While Cronqvist and Yu (2017) implicitly assume that corporate social responsibility efforts are the use of corporate resource not related to shareholders' interest, our study attempts to argue that sustainability efforts indeed contribute to shareholders' value.

high sustainability efforts show abnormal returns after the sustainability index listing.

It is investors' preferences that are propelling the trend in positively considering and identifying environmental, social, and governance practices, together with fundamental valuation metrics. Thus, it is interesting to ask whether such a sustainable focus yields financial benefits for investors over either short or long periods. However, the existing literature on stock performance is inconclusive.[2] In this paper, we reconsider this issue by examining the *long-term* stock performance of firms *first added* to the Dow Jones Sustainability Index (DJSI) North America. The DJSI North America is composed of the top 20 percent of the largest 600 stocks in the S&P Global Broad Market Index (BMI) in the USA and Canada based on their sustainability practices. The index selects companies "based on a comprehensive assessment of long-term economic, environmental, and social criteria that account for general as well as industry-specific sustainability trends (http://www.sustainability-indices.com)." Several management studies employ the DJSI to examine the effect of sustainability efforts on the firms' financial performance in various regions and different contexts (Consolandi et al. 2009; Doh et al. 2010; Cheung 2011; Robinson et al. 2011; Hawn et al. 2018). However, those studies mostly focus on the short-term effect of the DJSI membership.[3]

Our study is different from the existing management literature in that we attempt to provide an implication for investors by examining the *long-term* performance of sustainability stocks experiencing the DJSI listing event. Therefore, we employ the calendar-time abnormal return method to measure the long-term stock performance of firms first added to the DJSI. The calendar-time abnormal returns are the characteristics-adjusted monthly excess returns, developed by Daniel et al. (1997). This method is widely used to detect the long-term stock performance of firms experiencing a specific event such as managerial decisions (Mitchell and Stafford 2000) or insider trading (Jeng et al. 2003). Moreover, our event-study approach is different from the existing studies on sustainability stock performance in that we intend to capture the long-term performance of stocks *newly added* to the index, while the existing literature (i.e., Eccles et al. (2014)) estimates the buying-and-holding return of the *entire* sustainability portfolio formed periodically. We believe that this event-study approach allows us to clearly isolate the effect of the sustainability index listing on stock performance, compared with the zero-investment portfolio approach for the *entire* sustainability portfolio used in the existing research.

In our analysis, we find that stocks newly added to the sustainability index exhibit significantly positive abnormal returns for 12–30 months. Specifically, our finding is economically significant as the annualized abnormal return ranges from 4.85 to 3.64% (or 3.84–3.78%) against the value-weighted benchmark (or the equal-weighted benchmark) over the period ranging from 12 to 30 months after the index listing. However, we find no evidence that sustainability stocks generate significant excess returns for 11 months before firms are first added to the index, suggesting that the superior stock performance is triggered by the index listing. Therefore, the current empirical evidence is comparable to Eccles et al. (2014). Focusing on the organizational performance of sustainability companies, Eccles et al. (2014) provide evidence that high sustainability companies significantly outperform their counterparts in terms of stock market as well as accounting performance.

In the meantime, our analysis is a long-term event study for the sustainability index listing in that we attempt to confirm that the index listing is a value-enhancing event. Therefore, the stock price behavior after the index listing is similar to the "index recognition effect" which illustrates a positive stock price response to index listing (Shleifer 1986),[4] Beneish and Whaley (1996), Lynch and Mendenhall (1997), Wurgler and Zhuravskaya (2002), Denis et al. (2003), and Chen et al. (2004). However, other studies, including Harris and Gurel (1986), argue that an index inclusion only leads to a temporary price effect which should dissipate once the excess demand is satisfied. If sustainability efforts are indeed recognized through the index listing, stocks are likely to experience a significant price change in response to the sustainability index listing. Consistent with the index recognition effect, our analysis shows that the sustainability index listing leads to a long-term, positive price change, specifically a significant abnormal return over more than one year.[5]

Moreover, we find an asymmetric price behavior of sustainability stocks. Specifically, delisted stocks experience no

[2] Bonini and Swartz (2014) cite research by Deutsche Bank evaluating 56 academic studies that show companies with high ratings for environmental, social, and governance (ESG) factors have lower costs of debt and equity. However, Jones et al. (2007) report a generally negative relationship between sustainability disclosures of Australian firms and abnormal returns. Bianchi and Drew (2012) report that a recent performance of sustainable stock indices is worse than other indices over the long term, while De Haan et al. (2012) report a negative relationship between corporate environmental performance and stock returns.

[3] These studies use an event-study approach which maps stock market reactions to news regarding the DJSI membership based on the cumulative abnormal daily return method (CAR).

[4] Many studies address the permanent value effect attributable to index inclusion. See Shleifer (1986).

[5] We attempt to highlight the long-term performance of sustainability stocks based on the calendar-time monthly abnormal return method, while most of the literature on the index recognition effect focuses on the short-term price effect using the cumulative abnormal daily return method (CAR).

significant abnormal returns after the index delisting. That is, stocks newly added to the index show a permanent price increase, while delisted stocks exhibit no permanent price decline. This asymmetric price behavior is consistent with the permanent index effect. According to Chen et al. (2004), there is an asymmetric price response to changes to the S&P 500 index. They argue that investor awareness increases after addition to the index, but awareness does not easily diminish after deletion from the index.

In the following analysis, we examine more carefully what factor drives the price behavior of sustainability stocks. Our analysis suggests that actual actions corresponding to sustainability awareness and intentions translate into an increase in demand for sustainability stocks by investors, especially institutional investors. Specifically, we find that sustainability stocks exhibit a significant increase in institutional ownership after firms are first added to the sustainability index. The institutional ownership increases by 0.982–4.159% over the period of 12–30 months after the index listing. However, we find no significant drop in institutional ownership after the index delisting. This asymmetric response of institutional ownership is consistent with the price behavior of sustainability stocks. Moreover, it is also documented in the literature (Chan et al. 2013; Madhavan 2003; Cai and Houge 2008).

Finally, we conduct two additional analyses. On the one hand, we examine whether the sustainability index systematically picks up high-performing stocks, i.e., high-risk stocks which, in turn, are likely to generate high stock returns. On the other hand, we also investigate if sophisticated short sellers increase their position to exploit a possible overpricing after the index listing. In the analysis, we find no evidence that risk characteristics such as capitalization, book-to-market ratio, or volatilities contribute to the superior performance. However, the analysis shows that sustainability stocks exhibit a higher Sharpe ratio than non-sustainability, suggesting that sustainability stocks are superior performers in the mean–variance space. We also find that sustainability stocks experience no significant change in monthly short interest level around index listing/delisting. After all, the robustness analysis confirms that the abnormal performance is indeed driven by a growing demand for sustainability stocks and that the index recognition leads to a permanent price impact, not a temporary overpricing, consistent with our main theme.

The rest of the paper is organized as follows. In "Methodology and data" section, we describe the methodology and data. We discuss the main analysis in "Sustainability and stock performance" and "Sustainability stock performance and institutional ownership" sections. We provide the robustness check in "Discussions about risk characteristics and short sale activity" section. Finally, we conclude in "Conclusion" section.

Methodology and data

We examine the components of the Dow Jones Sustainability Index North America for each year from 2005 through 2016. This index is revised every September. It is a market capitalization weighted, broad common stock index. Companies are selected for the sustainability "indices based on a comprehensive assessment of long-term economic, environmental, and social criteria that account for general as well as industry-specific sustainability trends (http://www.sustainability-indices.com)." The index construction employs a rules-based methodology using primary research and focuses on best-in-class companies identified through SAM's extensive annual corporate sustainability assessment of ESG attributes such as corporate governance, climate strategy, tax strategy, and supply chain standards to establish a total sustainability score. The DJSI North America is composed of the top 20 percent of the largest 600 stocks in the S&P Global Broad Market Index (BMI) in the USA and Canada based on their sustainability practices.

We note the merit of this classification scheme. By using this methodology, we reduce the degree of subjectivity inherent in concluding that a firm is sustainability-centric or not. The lack of completely standardized and mandated accounting rules for disclosing "sustainability" makes using financial statement data imprecise in gauging a firm's commitment to sustainability-consistent activities and operations. Furthermore, sustainability is a long-term concept that makes the analysis of daily sustainability-related announcements not reflective of firms that have an established character of sustainability-consistent practices. For example, expenditures incurred on using more environmental efficient inputs are a cost in the short-run but may produce necessary innovations that show up in a firm over a longer period of time. A firm announcing a change in operations to sustain relationships with community stakeholders may not immediately translate into improved monetary performance. The difficulty in classifying what constitutes a sustainability announcement and the nature of sustainability as a long-term concept with long-term consequences are challenges to applying any methodology that attempts to investigate the impact of sustainability-like activities on firm performance. Our paper attempts to address these challenges by documenting the performance of firms that have been classified as sustainable-consistent in their operations by a third party. This mitigates the degree of subjectivity in differentiating between sustainability and non-sustainability-consistent firms.

Specifically, we focus on the performance of firms that are added to the index. As a result, the "sustainability"

Table 1 Construction of sustainability portfolio

Panel A: Construction of Sample			
			Total
# of firm-year observations			1460
Less: stocks not listed in the U. S.		– 32	
# of available firm-year observations			1428

Panel B: Number of Stocks

Year	Index	Delisted	Sample
2005	111		108
2006	113	16	17
2007	120	9	13
2008	124	14	13
2009	139	10	20
2010	136	19	15
2011	143	11	13
2012	140	14	10
2013	140	19	10
2014	149	9	9
2015	145	14	6
Total	1460	135	234

Panel C: Industry Classification

Industry	Sample
Consumer Nondurables/Durables	24
Manufacturing	26
Energy/Chemicals and Allied products	22
Business Equipment	28
Utilities/Telecom	27
Shops	24
Health	18
Money	33
Other	32
Total	234

The table presents the time-series average of characteristics for portfolios formed every September on the Dow Jones Sustainability Index North America. Sustainability portfolio consists of stocks listed for the first time on the index. The sample spans October 2005 to September 2016

portfolio consists of stocks listed on the index for the first time. The precise sample period is October 2005 to September 2016. (The last firms entering the sustainability portfolio are added in September 2015 so that there exist at least one year of performance data through September 2016.)

Table 1 summarizes the attributes of our sample. Panel A of Table 1 indicates that there are 1460 firm-year observations for the index during the period of 2005–2016. (Thirty-two Canadian firm-years are excluded since they are not listed on the US markets.) This results in 1428 firm-year observations. Panel B shows that 234 firms enter the sustainability portfolio. We focus only on common stocks (CRSP share code of 10 or 11). These are firms that are added at some point over the 2005–2016 period to the index. Half of sample stocks are gradually added to the sustainability portfolio over time. In contrast, total 135 firms drop out of the index over time. In Panel C, we show the distribution of firms over industries. We classify industry based on the Fama and French (1997) 12 industry classification using four digit SIC codes.

Table 2 presents the time-series average characteristics for the sustainability portfolio. Firm size is the logarithm of the market value of equity, measured in June of year t. Book-to-market is the logarithm of the book-to-market ratio of equity of fiscal year $t-1$. The book value of common equity is obtained from the annual Compustat files (Items 60 and 74). We use a firm's market equity in December of

Table 2 Descriptive statistics for sustainability portfolio

	Sample	Control	Difference
Firm size	9.672***	9.530***	0.142***
	0.030	0.028	0.012
Book-to-market	−0.910***	−0.930***	0.020
	0.031	0.025	0.014
Return on Asset	0.062***	0.061***	0.001
	0.001	0.002	0.002
Return on Equity	0.688***	0.186***	0.503***
	0.171	0.022	0.161
Capital Expenditure/Asset	0.046***	0.049***	−0.003***
	0.001	0.001	0.001
Cash Flow/Asset	0.098***	0.095***	0.003**
	0.001	0.002	0.001
Debt/Asset	0.255***	0.233***	0.022***
	0.006	0.002	0.005
Cash Balance/Asset	0.119***	0.134***	−0.016***
	0.003	0.002	0.004

The table presents the time-series average of characteristics for portfolios formed every September on the Dow Jones Sustainability Index North America. Sustainability portfolio consists of stocks listed for the first time on the index. A matching control portfolio is rebalanced monthly based on firm size, book-to-market, and momentum. Firm size is the logarithm of the market value of equity, measured in June of year t. Book- to-market is the logarithm of the book-to-market ratio of equity of fiscal year $t-1$. Momentum is the cumulative monthly return in $t-12$ through $t-1$. Return on asset is the ratio of net income to asset of fiscal year $t-1$. Return on equity is the ratio of net income to equity fiscal year $t-1$. Capital expenditures (Item 128) measure corporate investment activities. Cash flow is income before extraordinary items (Item 18) plus depreciation and amortization (Item 14). Debt is long-term debt (Item 9) plus debt in current liabilities (Item 34). Cash balance is cash and short-term investment (Item 1). All these items are scaled by total assets in fiscal year $t-1$. A standard error is reported below the average. The sample spans October 2005 to September 2016. Significance levels are as follows: 1% as "***," 5% as "**," and 10% as "*"

year $t-1$ to measure the book-to-market ratio. Return on asset is the ratio of net income (Item 177) to total asset (Item 6) of fiscal year $t-1$. Return on equity is the ratio of net income to the book value of common equity (Items 60 and 74) of fiscal year $t-1$. We also examine several balance sheet items to gauge a firm's financial status. Capital expenditures (Item 128) measure corporate investment activities. Cash flow is income before extraordinary items (Item 18) plus depreciation and amortization (Item 14). Debt is long-term debt (Item 9) plus debt in current liabilities (Item 34). Cash balance is cash and short-term investment (Item 1). All the balance sheet items are scaled by total assets as of fiscal year $t-1$.

We also construct the control portfolio. The matching control portfolio is rebalanced monthly based on firm size, book-to-market, and momentum. Momentum is the cumulative monthly return from $t-12$ through $t-1$. The following steps explain how the control portfolio is constructed. First, a universe of stocks is sorted based on firm size, book-to-market, and momentum. Both firm size and book-to-market are measured annually based on Fama and French (1993). Momentum is measured monthly. The sorted stocks are then classified into 25 groups (=5 size × 5 book-to-market ratio) annually. The breakpoints for firm size use all NYSE stocks that have a CRSP share code of 10 or 11 and have share and price data.[6] We exclude closed end funds and REITs. The breakpoints for the book-to-market ratios use all NYSE stocks for which we have the market value of equity (ME) for December of year $t-1$ and the positive book value of equity (BE) for the last fiscal year end in $t-1$. After the annual classification, all the stocks are reclassified monthly into 5 groups based on momentum. As a result, all the stocks are classified into 125 groups *monthly*. Therefore, each firm in our sample (234 firms) belongs to one of 125 groups. This classification is renewed every month. This group is subsequently used as the control portfolio. This control portfolio method is consistent with the calendar-time portfolio (CTP) approach used by Mitchell and Stafford (2000). This approach is designed to match sample firms and control firms based on characteristics determining stock performance.

Table 2 shows several notable characteristics for sustainability firms, compared with control firms. First, sustainability firms are larger than control firms even after controlling for firm size. However, the book-to-market ratio shows no significant difference between sustainability firms and control firms. Second, we see no significant difference in return on assets, while we find that sustainability firms show a higher return on equity than control firms. Third, we observe some evidence consistent with the extant literature. Sustainability firms show a higher level of operating cash flow and a lower level of cash balances than control firms, consistent with Jones et al. (2007). Interestingly, we find that sustainability firms use significantly higher financial leverage than control firms, as shown in debt-to-total asset ratios.

In Table 3, we present the average of characteristics for sustainability firms across industries. We follow the Fama and French (1997) 12 industry classification. Table 3 highlights distinctive characteristics for sustainability firms within their industry group, while Table 2 examines sustainability stocks from the asset pricing point of view. We find several interesting attributes of sustainability firms, compared with their industry peers. First, sustainability firms are larger than their industry peer group, while their book-to-market ratio is lower than their peer ratio. Both

[6] NYSE breakpoints data for firm size and book-to-market ratio were obtained from the Fama–French website at: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Breakpoints.

Table 3 Sustainability portfolio characteristics across industry

	Firm size			Book-to-market		
	Sample	Average	Difference	Sample	Average	Difference
Nondurables/Durables Non-durables/Durables	9.750	6.217	3.533***	−1.138	−0.666	−0.472***
Manufacturing	9.493	6.280	3.213***	−1.163	−0.735	−0.427***
Energy/Chemicals Allied products	10.236	6.853	3.383***	−1.009	−0.738	−0.271*
Business Equipment	10.000	5.774	4.226***	−1.261	−0.925	−0.337***
Utilities/Telecom	9.472	6.893	2.579***	−0.167	−0.344	0.178***
Shops	9.668	6.053	3.615***	−1.036	−0.609	−0.427***
Health	10.299	5.667	4.632***	−1.392	−1.207	−0.185*
Money	10.197	5.867	4.330***	−0.692	−0.438	−0.254***
Other	9.554	5.931	3.632***	−0.879	−0.767	−0.115
Total	9.836	6.161	3.675***	−0.942	−0.692	−0.251***
	Return on asset			Return on equity		
	Sample	Average	Difference	Sample	Average	Difference
Nondurables/Durables Non-durables/Durables	0.084	0.008	0.077***	0.247	−0.139	0.386***
Manufacturing	0.069	0.018	0.051***	2.879[a]	0.027	2.852
Energy/Chemicals Allied products	0.082	0.013	0.069***	0.181	0.000	0.181***
Business Equipment	0.078	−0.034	0.112***	0.141	−0.186	0.327***
Utilities/Telecom	0.018	−0.010	0.028***	0.039	−0.206	0.245***
Shops	0.083	0.020	0.063***	0.182	−0.037	0.219***
Health	0.101	−0.187	0.288***	0.182	−0.679	0.861***
Money	0.021	0.011	0.009	0.144	0.018	0.103***
Other	0.064	−0.055	0.117***	0.145	−0.149	0.293***
Total	0.067	−0.019	0.088***	0.484	−0.131	0.625*
	Capital expenditure/asset			Cash flow/asset		
	Sample	Average	Difference	Sample	Average	Difference
Nondurables/Durables Non-durables/Durables	0.037	0.042	−0.005	0.118	0.073	0.045***
Manufacturing	0.041	0.044	−0.003	0.112	0.073	0.040***
Energy/Chemicals Allied products	0.107	0.176	−0.070***	0.164	0.098	0.066***
Business Equipment	0.052	0.038	0.014	0.147	0.029	0.118***
Utilities/Telecom	0.058	0.067	−0.010***	0.069	0.045	0.023***
Shops	0.071	0.066	0.005	0.142	0.084	0.058***
Health	0.037	0.041	−0.003	0.153	−0.135	0.288***
Money	0.009	0.016	−0.006	0.027	0.017	0.011
Other	0.067	0.162	−0.095**	0.119	−0.084	0.202***
Total	0.055	0.073	−0.021***	0.113	0.023	0.090***
	Debt/asset			Cash balance/asset		
	Sample	Average	Difference	Sample	Average	Difference
Nondurables/Durables Non-durables/Durables	0.329	0.235	0.093***	0.104	0.158	−0.055***
Manufacturing	0.259	0.227	0.031	0.134	0.149	−0.015
Energy/Chemicals Allied products	0.345	0.292	0.053	0.079	0.725	−0.646*

Table 3 (continued)

	Debt/asset			Cash balance/asset		
	Sample	Average	Difference	Sample	Average	Difference
Business Equipment	0.139	0.125	0.014	0.293	0.413	−0.121***
Utilities/Telecom	0.385	0.362	0.023	0.028	0.123	−0.096***
Shops	0.185	0.215	−0.030	0.153	0.146	0.006
Health	0.251	0.164	0.087***	0.199	0.545	−0.346***
Money	0.223	0.292	−0.070	0.236	1.816	−1.580***
Other	0.277	0.823	−0.564	0.121	0.644	−0.531***
Total	0.264	0.322	−0.059	0.146	0.566	−0.420***

The table presents the average of characteristics for portfolios formed every September on the Dow Jones Sustainability Index North America across industry based on the Fama and French (1997) 12 industry classification. Sustainability portfolio consists of stocks listed for the first time on the index. Firm size is the logarithm of the market value of equity, measured in June of year t. Book-to-market is the logarithm of the book-to-market ratio of equity of fiscal year $t-1$. Return on asset is the ratio of net income to asset of fiscal year $t-1$. Return on equity is the ratio of net income to equity of fiscal year $t-1$. Capital expenditures (Item 128) measure corporate investment activities. Cash flow is income before extraordinary items (Item 18) plus depreciation and amortization (Item 14). Debt is long-term debt (Item 9) plus debt in current liabilities (Item 34). Cash balance is cash and short-term investment (Item 1). All these items are scaled by total assets in fiscal year $t-1$. The sample spans October 2005 to September 2016. Significance levels are as follows: 1% as "***," 5% as "**," and 10% as "*"

[a]The number is due to the fact that Lockheed Martin shows a dramatic decrease in its equity in the fiscal year of 2012. In 2017, the common equity capital for Lockheed Martin is even negative

characteristics imply that sustainability stocks are not likely to be high-performing stocks associated with the small firm effect or the high book-to-market ratio effect. Second, sustainability firms generate more profits than the industry average based on both return on asset and return on equity. Third, sustainability firms exhibit a higher level of cash flow and a higher financial leverage than the industry average, while they maintain a lower level of cash balances, consistent with Table 2.

Sustainability and stock performance

In the main analysis, we examine the market reaction to, especially stock performance of, firms that have been first added to the sustainability index. In particular, we hypothesize that the sustainability index recognition is a value-enhancing event, leading to an increase in stock price for those firms. The sustainability index recognition raises the awareness of a firm which proactively takes into account sustainability issues in its managerial decision to meet the needs of potential investors and customers. Therefore, sustainability stocks can attract investors who positively consider environmental, social, and governance practices together with fundamental valuation metrics. As a result, those stocks are likely to show a positive price response to the index listing.

To address this issue, we employ the method widely accepted and designed to capture stock performance following an event, the calendar-time abnormal return method. The calendar-time abnormal returns are the characteristics-adjusted monthly excess returns, developed by Daniel et al. (1997). Mitchell and Stafford (2000) and Jeng et al. (2003) use this method to estimate the long-term stock performance of managerial decisions or inside trading. The return method first subtracts the value or equal-weighted return of the control portfolio from the return of sustainability stocks to estimate the characteristics-adjusted monthly excess returns for individual sustainability stocks. The average monthly excess return for the sustainability portfolio is then computed for month t. Following this step, we calculate the time-series average of the characteristic-adjusted monthly excess return for the sustainability portfolio over a J month holding period. This return adjustment is designed to control stylized factors such as firm size, book-to-market, and momentum in assessing stock performance.

Table 4 presents the calendar-time abnormal returns for the sustainability portfolio. Panel A of Table 4 reports the analysis for stocks newly added to the sustainability index. Consistent with the hypothesis, Panel A of Table 4 shows that stocks newly added to the index exhibit positive and significant abnormal returns for 12–30 months at 5% one-sided significance level after firms are first added to the sustainability index. The finding is economically significant as the annualized abnormal return ranges from 4.85 to 3.64% (or 3.84–3.78%) against the value-weighted benchmark (or the equal-weighted benchmark) over the period ranging from 12 to 30 months after the sustainability index listing. Beyond 30 months, we observe weaker or insignificant excess returns for stocks newly added to the sustainability index. Despite the short sample period, this analysis suggests that the sustainability index listing is indeed a value-enhancing event, leading to a persistent increase in stock price for those firms.

Table 4 Calendar-time abnormal returns for stocks after sustainability listing/delisting

	Panel A: Stocks added to the index for the first time and held for J-months								
	$J = -11$	−6	0	6	12	18	24	30	36
Value-weighted benchmark	0.091	0.051	−0.484	0.273	0.404	0.416	0.410	0.303	0.198
	0.43	0.18	−0.65	0.89	1.77*	2.00**	2.06**	1.73*	1.41
Equal-weighted benchmark	0.093	0.065	−0.036	0.205	0.320	0.358	0.407	0.315	0.212
	0.57	0.32	−0.09	0.77	1.74*	2.14**	2.41***	2.16**	1.93*
	Panel B: Stocks dropped from the index for the first time and held for J-months								
	$J = -11$	−6	0	6	12	18	24	30	36
Value-weighted benchmark	0.215	0.077	−2.156	−0.043	0.226	0.113	−0.130	−0.206	−0.198
	0.80	0.19	−1.72*	−0.11	0.57	0.30	−0.52	−0.91	−1.10
Equal-weighted benchmark	0.075	−0.056	−1.320	0.168	0.363	0.219	−0.027	−0.085	−0.107
	0.33	−0.17	−1.53	0.47	1.00	0.63	−0.11	−0.39	−0.60

The table presents the time-series average of the Daniel et al. (1997) characteristic-adjusted monthly excess returns over the period of *J*-months after stocks are added to/dropped from the Dow Jones Sustainability North America Index every September. A matching benchmark portfolio is rebalanced monthly based on firm' size, book-to-market, and momentum. Firm size is the logarithm of the market value of equity, measured in June of year *t*. Book-to-market is the logarithm of the book-to-market ratio of equity of fiscal year $t-1$. Momentum is the cumulative monthly return in $t-12$ through $t-1$. A *t*-statistic is reported below returns. The sample spans October 2005 to September 2016. Significance levels are as follows: 1% as "***," 5% as "**," and 10% as "*"

Moreover, this analysis also confirms the existing study addressing the positive effect of sustainability efforts on the market value of sustainability firms, such as Eccles et al. (2014).

In Panel A of Table 4, we also report the performance of sustainability stocks prior to the index listing to check whether sustainability stocks persistently outperform non-sustainability stocks even before they join the index. However, the analysis shows that sustainability stocks generate no significant excess returns for 11 months before firms are first added to the sustainability index. In other words, this analysis suggests that the superior performance of sustainability stocks is triggered by the index recognition event, not due to any characteristics associated with sustainability stocks.

Panel B of Table 4 shows the reaction of stock price to the index delisting. According to Table 1, the index drops total 135 firms over time. If the sustainability index listing is a value-enhancing event, the index delisting might result in a decrease in stock price or at least no excess returns for those firms. Therefore, we investigate this possibility by estimating excess returns after firms are first delisted from the sustainability index. Panel B of Table 4 shows that delisted stocks show no excess returns after firms are excluded from the index, while they experience an insignificant stock price drop during the revision period, as shown in month = 0. Interestingly, delisted stocks exhibit no excess returns over the 11-month period even before the index delisting. In fact, this asymmetric behavior of stock price is consistent with the literature on the index recognition effect. According to Chen et al. (2004), there is an asymmetric price response to changes to the S&P 500 index. That is, stocks newly added to the S&P 500 index show a permanent price increase, while deleted stocks exhibit no permanent price decline. A possible explanation is based on the change in investor awareness. They argue that investor awareness permanently increases after stocks join the index that investors pay attention to, but their awareness does not easily diminish even after those stocks are excluded from the index due to any reason. Therefore, as long as the sustainability index listing is a permanent value-enhancing event, the index delisting does not devalue sustainability stocks due to an increase in investor awareness, resulting in no significant change in stock price for those firms.

Overall, the evidence shows that an integrated selection procedure that augments financial considerations with high sustainability standards produces superior investment returns. In particular, the sustainability index listing is a value-enhancing event and leads to a permanent increase in stock price by raising the interest of investors in sustainability firms. Therefore, our analysis is consistent, in spirit, with the literature on the index recognition effect which illustrates a positive stock price response to the index listing (Shleifer 1986). In addition to Chen et al. (2004), Chan et al. (2013) consider the impact of inclusion/exclusion in the S&P 500 on stock returns, while both Madhavan (2003) and Cai and Houge (2008) analyze the effect of the Russell index addition/deletion. However, our study is different from the existing studies on the index effect in that we analyze the relatively long-run performance of stocks following their index inclusion by employing the calendar-time abnormal return method. Moreover, our analysis also confirms the permanent value effect of the index listing, compared with the temporary price effect proposed by Harris and Gurel (1986).

Sustainability stock performance and institutional ownership

In the following analysis, we investigate what economic factor drives the superior performance of sustainability stocks. Specifically, we examine whether the abnormal performance is driven by a growing interest in sustainability stocks. In fact, the current market trend supports this hypothesis. Recent trends point to a growing interest in sustainability stocks as an investment allocation. For example, according to Bloomberg, "mainstream investment firms are rushing into 'sustainable investing,' also known as SRI (socially responsible investing) and ESG (environment, social, and corporate governance)." In the period from 2012 to 2014, the number of US investment funds integrating ESG criteria increased by 28 percent and their assets more than quadrupled to $4.3 trillion. BlackRock, the largest money manager in the world, started its sustainability-oriented BlackRock Impact US Equity Fund on October 13, 2015. (https://www.bloomberg.com/news/articlesESGcriteriaincreased/2015-10-21/sustainable-investing-is-booming-is-it-smart-). In the meantime, the integration of sustainability and ESG criteria into investment management practice and education has been validated by the CFA Institute. The following quote from the CFA Institute, Environmental, Social and Governance Issues in Investing: A Guide for Investment Professionals, 2015 reflects this emphasis: "For investment professionals, a key idea in the discussion of ESG issues is that systematically considering ESG issues will likely lead to more complete investment analyses and better-informed investment decisions." Considering this trend, we conjecture that the sustainability index listing leads to an increase in investor interest in those stocks, resulting in an increase in demand by institutional investors after firms are newly added to the sustainability index.

According to the literature (Madhavan 2003), there are two possible explanations for the index recognition effect: the liquidity approach and the value approach. On the one hand, the liquidity approach argues that transitory order imbalances associated with index additions and deletions are the primary source of price movements. So, the price effect will disappear once excess demand is satisfied. On the other hand, the value approach suggests that the index itself is a source of value, possibly because of changes in information flows or liquidity. Therefore, the price effect will be permanent even after stocks lose the index membership. Therefore, given the continuing trend of incorporating sustainability criteria with conventional financial metrics, we hypothesize that sustainability awareness leads to a permanent increase in demand for sustainability stocks by investors, especially institutional investors. As the sustainability index recognition enhances firm value, we expect to observe a significant increase in institutional ownership after stocks are newly added to the sustainability index, but we anticipate no significant change in institutional ownership after the index deletion, consistent with the asymmetric price behavior of sustainability stocks as shown in the previous section.

The following illustrates the method to measure a change in institutional ownership of sustainability stocks after the index listing/delisting. First, we obtain the September institutional ownership data from the 13F filings in the CDA Spectrum database to measure a baseline ownership because the sustainability index is revised every September. Second, we calculate the average percentage change in institutional ownership of sustainability stocks over the period of 3–36 months after the index listing. Third, we also examine the same metric for stocks delisted from the index to illustrate the reaction of institutional investors to the event which is similar in nature but opposite in direction to the index listing event. Table 5 presents the analysis.

Panel A of Table 5 presents the percentage change in institutional ownership after firms are newly added to the sustainability index. The analysis suggests that actual actions corresponding to sustainability awareness lead to an increase in demand for sustainability stocks by institutional investors, supporting our hypothesis. Specifically, we find that sustainability stocks exhibit a significant increase in institutional ownership after the index listing. The institutional ownership significantly increases by 0.982% for 12 months. Moreover, we observe a persistent increase in institutional ownership over the period of 24 to 30 months, 2.740 to 4.159%.[7] A continuous presence in the index for some sustainability stocks can contribute to the persistence in the ownership trend beyond annual index revision.[8]

Panel B of Table 5 shows the change in institutional ownership after the index delisting. We find no significant change in institutional ownership for sustainability stocks after the index delisting. Specifically, the institutional ownership decreases by an insignificant degree of −0.887% for the first 12 months after the index delisting. Over the subsequent period, we observe a similar insignificant change in institutional ownership of 0.010–0.758% for 24–36 months. This pattern is in a striking contrast to a change in institutional ownership in response to the index listing for sustainability

[7] We still obtain a similar but weak result when we separate the initial listing of 108 firms from the subsequent listing of 126 firms. The result is available upon request.

[8] A secular trend in institutional ownership can partially explain an increase in the ownership for sustainability stocks. However, this explanation is inconsistent with the analysis for delisting stocks in Panel B of Table 5.

Table 5 Change in institutional ownership after sustainability listing/delisting

	Panel A: Sustainability Index Listing							
	$J=3$	6	9	12	18	24	30	36
Percentage Change (%)	0.541	0.915	1.253	0.982	2.368	2.740	4.159	1.700
	1.06	2.20**	3.13***	2.32***	4.46***	3.64***	5.70***	2.82***
	Panel B: Sustainability Index Delisting							
	$J=3$	6	9	12	18	24	30	36
Percentage Change (%)	0.062	0.155	0.704	−0.887	0.289	0.010	−0.911	0.758
	0.11	0.23	0.95	−0.82	0.41	0.01	−0.81	0.88

The table presents the percentage change in institutional ownership over the period of *J*-months after stocks are listed on or delisted from the Dow Jones Sustainability Index North America. The baseline ownership is calculated in September because the Dow Jones Sustainability North America Index is revised every September. A *t*-statistic is reported below. The sample spans October 2005 to September 2016. Significance levels are as follows: 1% as "***," 5% as "**," and 10% as "*"

stocks, as shown in Panel A of Table 5. Therefore, this evidence is consistent with the hypothesis. Moreover, this asymmetric behavior of institutional investors is well documented in the permanent index effect associated with many indices. Chan et al. (2013) address that addition to the S&P 500 index leads to a permanent increase in institutional ownership, while deletion has no significant effect on the institutional ownership.

Overall, we find that the sustainability index listing boosts the demand for stocks which proactively take into account sustainability issues in their managerial decision to meet the needs of potential investors, especially of institutional investors who value sustainability awareness and efforts. Moreover, we find no significant decrease in institutional ownership after the index delisting. After all, the current analysis provides a nice explanation for the empirical evidence that sustainability stocks exhibit a superior performance over the period of at least 30 months after the index listing and experience no significant price drop even after the index delisting, supporting the asymmetric stock price response to changes to the index membership, as shown in the previous section.

Discussions about risk characteristics and short sale activity

In the next analysis, we investigate two interesting issues. First, we examine whether the superior performance of sustainability stocks is driven by any risk characteristics. Second, we also attempt to analyze the behavior of sophisticated investors, especially short sellers, around sustainability index listing/delisting.

The intuition for the first analysis is as follows. If the superior performance of sustainability stocks was due to their risk profile, the positive abnormal return would be attributable to their higher risk exposure, not to an increase in demand for sustainability stocks by investors. So, we investigate the risk-based alternative explanation. In particular, we compare sustainability stocks with the control portfolio of non-sustainability stocks by analyzing four volatility measures: daily return volatility, CAPM beta, and two idiosyncratic volatilities based on CAPM and the Fama and French (1993) three-factor model, respectively. In addition, we also measure the Sharpe ratio to examine the risk-adjusted performance of sustainability stocks compared with non-sustainability stocks.[9]

Table 6 summarizes the empirical analysis. Overall, the analysis shows that the superior performance of sustainability stocks is not attributable to their higher risk exposure. In other words, sustainability stocks are characterized by lower, not higher, risk than their counterparts. Specifically, daily return volatility and CAPM beta are 0.017 and 1.004 for the sustainability portfolio, while the same metrics are 0.032 and 1.398 for the control portfolio. Similarly, the CAPM and Fama–French idiosyncratic volatilities are 0.076 and 0.074 for sustainability stocks, while they are 0.128 and 0.126 for the non-sustainability counterparts, respectively. Moreover, this evidence is also consistent with characteristics of sustainability stocks, as shown in Tables 2 and 3. According to the descriptive statistics in Tables 2 and 3, sustainability firms are larger than non-sustainability stocks, while their book-to-market ratio is lower than the industry average. In the meantime, the analysis also shows that sustainability stocks exhibit a higher Sharpe ratio than non-sustainability stocks, as shown in the last column of Table 6. This suggests that sustainability stocks are superior performers in the mean–variance space where we consider reward and risk

[9] We would like to thank the referee for suggesting this analysis.

Table 6 Volatility and Sharpe ratio analysis for sustainability stocks

	Volatility for sustainability/non-sustainability stocks				
	Daily return volatility	CAPM beta	CAPM idiosyncratic risk	Fama–French idiosyncratic risk	Sharpe ratio
Sustainability stocks	0.017	1.004	0.076	0.074	0.128
Non-sustainability stocks	0.032	1.398	0.128	0.126	0.087

The table presents volatility and Sharpe ratio. Volatility measures include the average percentage daily return volatility over prior year, the average CAPM beta, the average percentage residual volatility from regressing 60-month excess returns on the market excess return or the Fama and French (1993) three factors: market excess return ($rm-rf$), small stock returns minus large stock returns (SMB), and high book-to-market stock returns minus low book–to–market stock returns (HML). rp, $t-rf$, $t=\alpha+\beta 1$ (rm, $t-rf$, $t)+\beta 2$ SMB$+\beta 3$ HML$+\varepsilon$ p, t Sustainability stocks are selected September on the Dow Jones Sustainability North America Index. The sample spans October 2005 to September 2016

Table 7 Change in short interest after sustainability listing/delisting

	Panel A: Sustainability Index Listing									
	$J=1$	2	3	6	9	12	18	24	30	36
Percentage Change (%)	−0.001	−0.001	−0.001	−0.001	0.001	0.001	0.000	0.001	0.008	0.012
	−1.17	−0.90	−1.37	−1.01	0.64	0.55	0.19	0.65	2.74***	3.38***
	Panel B: Sustainability Index Delisting									
	$J=1$	2	3	6	9	12	18	24	30	36
Percentage Change (%)	−0.002	−0.003	−0.001	0.001	0.001	0.002	0.005	0.010	0.008	0.009
	−1.23	−1.65*	−0.58	0.41	0.41	0.66	0.96	1.73*	1.36	1.90*

The table presents the percentage change in short interest over the period of J-months after stocks are listed on or delisted from the Dow Jones Sustainability Index North America. Short interest is defined as a number of shares shorted to a number of outstanding shares in month t. The baseline short interest is calculated in September because the Dow Jones Sustainability North America Index is revised every September. A t-statistic is reported below. The sample spans October 2005 to September 2016. Significance levels are as follows: 1% as "***," 5% as "**," and 10% as "*"

simultaneously. Taken together, sustainability stocks are not characterized with a high-risk profile but with a high potential for risk-adjusted performance.

The next analysis focuses on the reaction of smart money or the behavior of sophisticated investors when stocks join the sustainability index or when stocks drop out of the index. Suppose that the market overreacted to the sustainability index inclusion of stocks and thus sustainability stocks were temporarily overpriced. Then, we would expect that sophisticated arbitragers exploit the arbitrage opportunity by shorting sustainability stocks after the index listing. In other words, we explore the market overreaction hypothesis. We investigate this issue by examining monthly short interest, defined as the number of shares short relative to a number of outstanding shares. In particular, we measure a baseline monthly short interest of sustainability stocks in September and calculate the average change in short interest over the period of 1–36 months after the index listing, similar to the analysis for institutional ownership change. We also examine the same metric for stocks delisted from the sustainability index.

Table 7 presents the analysis. Overall, the analysis shows no significant change in short interest for sustainability stocks after the index listing or delisting. In other words, sophisticated short sellers do not change their position in response to the index membership change. On one hand, sustainability stocks experience no significant increase in short sale activity after they join the index listing. For example, short interest insignificantly fluctuates with a decrease of 0.001% for 1–6 months and an increase of 0.001% for 12–24 months, respectively, as shown in Panel A of Table 7. This short sale trend is in contrast to the abnormal performance of sustainability stocks over the same period after the index listing, as shown in Panel A of Table 4. On the other hand, we find a small but insignificant decrease in short interest for 1–3 months after stocks are delisted from the sustainability index, as shown in Panel B of Table 7. Taken together, our analysis for short interest suggests that short sellers are not interested in the sustainability index change or they take no arbitrage action to exploit a temporary overpricing because they agree that the index recognition has a permanent impact on price, consistent with our main theme.

In summary, our risk analysis concludes that the superior performance of sustainability firms is not attributable to any risk characteristics associated with those stocks. However, sustainability stocks are superior performers in the mean–variance space, suggesting that sustainability stocks are not characterized with a high-risk profile but with a high potential for risk-adjusted performance. Moreover, our short sale analysis implies that the sustainability index listing is not consistent with the market overreaction hypothesis.

Conclusion

We find that a persistent dedication to meeting challenging sustainability criteria does not hinder stock performance. Specifically, we suggest that the sustainability index listing boosts the demand of institutional investors who value sustainability awareness and efforts, significantly contributing to the abnormal performance of sustainability stocks. We also confirm that the superior performance of firms satisfying the extensive and detailed screening required for DJSI inclusion is not attributable to any systematic factor, such as high volatility or high risk characteristics related to the size effect and the book-to-market effect. Lastly, we find no evidence that sophisticated short sellers take an arbitrage action to exploit a possible overpricing after the index listing.

This analysis clearly has implications for corporate managers as well as for investors. For managers, the additional costs of enhancing firm sustainability efforts are not detrimental to firm performance. For investors, a holistic approach that considers sustainability and ESG criteria with conventional financial metrics produces enhanced investment performance. This analysis is also of practical significance and points to future directions for sustainability and ESG integration research. The framework employed by SAM in conducting their corporate sustainability assessment includes a Media and Stakeholder Analysis (MSA) that is a structured, comprehensive, and apparently effective approach in identifying improved risk-return investment opportunities.

In sum, our paper contributes to the literature documenting the connection between firm sustainability efforts and stock performance by injecting some empirical insights into the literature.

References

Albuquerque, R., Y. Koskinen, and C. Zhang. 2019. Corporate social responsibility and firm risk: theory and empirical evidence. *Management Science* 65: 4451–4469.

Beneish, M., and R.E. Whaley. 1996. An anatomy of the "S&P 500 Game": the effects of changing the rules. *Journal of Finance* 51: 1909–1930.

Bianchi, R.J., and E. Drew. 2012. Sustainable stock indices and long-term portfolio decisions. *Journal of Sustainable Finance and Investment* 2: 303–317.

Bonini, S., and S. Swartz. 2014. *Profits with purpose: How organizing for sustainability can benefit the bottom line?*. New York: McKinsey and Company.

Cai, J., and T. Houge. 2008. Long-term impact of Russell 2000 index rebalancing. *Financial Analysts Journal* 65: 76–91.

Chan, K., H.W. Kot, and G.Y.N. Tang. 2013. A comprehensive long-term analysis of S&P 500 index additions and deletions. *Journal of Banking and Finance* 37: 4920–4930.

Chen, H., G. Noronha, and V. Singal. 2004. The price response to S&P 500 index additions and deletions: evidence of asymmetry and a new explanation. *Journal of Finance* 59: 1901–1930.

Cheng, I., H. Hong, and K. Shue. 2016. *Do managers do good with other peoples' money?* Working paper, Dartmouth College, Princeton University, and University of Chicago.

Cheung, A. 2011. Do stock investors value corporate sustainability? Evidence from an event study. *Journal of Business Ethics* 99: 145–165.

Consolandi, C., A. Jaiswal-Dale, E. Poggiani, and A. Vercelli. 2009. Global standards and ethical stock indexes: the case of the DJSI. *Journal of Business Ethics* 87: 185–197.

Cronqvist, H., and F. Yu. 2017. Shaped by their daughters: executives, female socialization, and corporate social responsibility. *Journal of Financial Economics* 126: 543–562.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers. 1997. Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance* 52: 1035–1058.

De Haan, M., L. Dam, and B. Scholtens. 2012. The drivers of the relationship between corporate environmental performance and stock market returns. *Journal of Sustainable Finance and Investment* 2: 338–375.

Denis, D., J. McConnell, A. Ovtchinnikov, and Yu. Yun. 2003. S&P 500 index additions and earnings expectations. *Journal of Finance* 58: 1821–1840.

Doh, J.P., S.D. Howton, S.W. Howton, and D.S. Siegel. 2010. Does the market respond to an endorsement of social responsibility? The role of institutions, information and legitimacy. *Journal of Management* 36 (6): 1461–1485.

Dow Jones Sustainability Index North America constituent member data. 2019. https://www.robecosam.com/csa/csa-resources/dow-jones-sustainability-indices-components.html.

Eccles, R.G., I. Ioannou, and G. Serafeim. 2014. The impact of corporate sustainability on organizational processes and performance. *Management Science* 60: 2835–2857.

Fama, E., and K. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33: 3–56.

Fama, E., and K. French. 1997. Industry cost of equity. *Journal of Financial Economics* 43: 153–193.

Harris, L., and E. Gurel. 1986. Price and volume effects associated with changes in the S&P 500: new evidence for the existence of price pressures. *Journal of Finance* 41: 815–830.

Hawn, O., K. Aaron, and W.M. Chatterji. 2018. Do investors actually value sustainability? New evidence from investor reactions to the Dow Jones Sustainability Index (DJSI). *Strategic Management Journal* 39 (4): 949–976.

Hong, H.G. and I. Liskovich. 2016. *Crime, punishment, and the value of corporate social responsibility*. Working Paper, Princeton University.

Jeng, L.A., A. Metrick, and R. Zeckhauser. 2003. Estimating the returns to insider trading: a performance evaluation perspective. *Review of Economics and Statistics* 85: 453–471.

Jones, S., G. Frost, J. Loftus, and S. Laan. 2007. An empirical examination of the market returns and financial performance of

entities engaged in sustainability reporting. *Australian Accounting Review* 17: 78–87.
Kitzmueller, M., and J. Shimshack. 2012. Economic perspectives on corporate social responsibility. *Journal of Economic Literature* 50: 51–84.
Krosinsky, C., N. Robins, and V. Stephen. 2011. *Evolutions in sustainable investing: strategies, funds and thought leadership*. Hoboken: Wiley.
Lins, K., H. Servaes, and A. Tamayo. 2017. Social capital, trust, and firm performance: the value of corporate social responsibility during the financial crisis. *Journal of Finance* 72: 1785–1824.
Lynch, A.W., and R.R. Mendenhall. 1997. New evidence on stock price effects associated with changes in the S&P 500 index. *Journal of Business* 70: 351–383.
Madhavan, A. 2003. The Russell reconstitution effect. *Financial Analysts Journal* 59: 51–64.
Masulis, R.W., and S.W. Reza. 2015. Agency problems of corporate philanthropy. *Review of Financial Studies* 28: 592–636.
McKinsey and Company. 2014. *Sustainability's strategic worth: McKinsey Global Survey results.* http://www.mckinsey.com/business-functions/sustainability-and-resource-productivity/our-insights/sustainabilitys-strategic-worth-mckinsey-global-survey-results.
Mitchell, M.L., and E. Stafford. 2000. Managerial decisions and long-term stock price performance. *Journal of Business* 73: 287–329.
RobecoSAM AG, Josefstrasse 218, 8005 Zurich Switzerland.
Robinson, M., A. Kleffner, and S. Bertels. 2011. Signaling sustainability leadership: empirical evidence of the value of DJSI membership. *Journal of Business Ethics* 101 (3): 493–505.
Servaes, H., and A. Tamayo. 2013. The impact of corporate social responsibility on the value of the firm: the role of customer awareness. *Management Science* 59: 1045–1061.
Shleifer, A. 1986. Do demand curves for stocks slope down? *Journal of Finance* 41: 579–590.
Wurgler, J., and K. Zhuravskaya. 2002. Does arbitrage flatten demand curves for stocks? *Journal of Business* 75: 583–608.

Moonsoo Kang is an Assistant Professor at Brooklyn College of the City University of New York.

K. G. Viswanathan is an Associate Professor at Hofstra University.

Nancy A. White is an Associate Professor at Hofstra University.

Edward J. Zychowicz is a Professor at Hofstra University.

Journal of Asset Management (2021) 22:133–150
https://doi.org/10.1057/s41260-020-00203-z

ORIGINAL ARTICLE

Expected and realized returns on stocks with high- and low-ESG exposure

Olaf Stotz[1]

Revised: 12 December 2020 / Accepted: 28 December 2020 / Published online: 12 January 2021

Abstract
Empirically, stocks with a good environmental, social, or governance (ESG) rating tend to earn higher returns than stocks with a low rating. In contrast, the expected returns of high-ESG stocks are primarily lower than those of low-ESG stocks. The difference between realized and expected returns in the ESG domain constitutes a puzzle which we will address in this paper. Applying a return decomposition, we find that the puzzle can be explained by discount rate news. We find that discount rates of high-ESG stocks have fallen relative to low-ESG stocks. However, discount rate news does not reflect changes in risk; rather, discount rate news is systematically related to the demand of investors who have ESG preferences.

Keywords Expected returns · Realized returns · Cash-flow news · Discount rate news · ESG · Sustainable investing

JEL Classification G12 · G30

Introduction

Socially responsible investing (SRI) has grown substantially over the recent years. According to the Financial Times (FT 2018), assets under management in funds that use environmental, social, or governance (ESG) screens have grown more than 600 percent to $23tn in the ten years to the end of 2016. Even conventional asset managers nowadays pay attention to ESG information (e.g., van Duuren et al. 2016). The general trend toward ESG investing shows that investor demand may potentially be driven by non-financial issues, such as social and ecological characteristics. This view is in contrast to the classical approach of asset management, which relies on the assumption that only financial issues, such as risk and return, are considered. In their theoretical approach, Fama and French (2007) describe how investor demand arising from non-financial factors may affect asset prices. Using their model's implications, we investigate the relation between ESG scores and stock returns from a novel perspective, which we call *the ESG return puzzle*. This phenomenon refers to the observation that *realized* stock returns tend to be *positively* related to ESG, while *expected* returns tend to be *negatively* related to ESG.

In efficient capital markets, return realizations should equal their expectations in the long-run (e.g., Fama 1991). In limited sample periods, however, a deviation of realized returns from expected returns (i.e., an unexpected return) can be explained by unexpected news (e.g., Campbell and Shiller 1988). In the context of the ESG return puzzle (unexpected returns tend to be positive), two potential news channels should be observed. First, cash-flow news suggests that high-ESG stocks should deliver positive surprises about future cash flows relative to low-ESG stocks. It should be noted that it is not the level of future cash-flow growth but expectations about how growth will change that drive the unexpected returns. There is theoretical support for such a cash-flow channel from the stakeholder theory approach (e.g., Jensen 2002). For example, customers may be more loyal to high-ESG companies and potentially pay higher prices. If the ESG preferences of customers increase over time, more individuals may pay higher prices, and this mechanism will lead to an upward revision of cash-flow expectations of a firm. Second, unexpected returns can also

Larry Fink (CEO Blackrock): "Sustainable investing will be a core component for how everyone invests. We are only at the early stages." FTfm, 19 Nov. 2018, page 6.

Olaf Stotz
o.stotz@fs.de

[1] Finance Department, Frankfurt School of Finance and Management, Adickesallee 32-34, 60322 Frankfurt, Germany

be driven by discount rate news. High-ESG stocks could realize higher returns than expected if investors apply an unexpectedly decreasing discount rate. Such a decrease in the discount rate can be attributed to two different mechanisms. The first of these relates changes in the discount rate to changes in the risk characteristics of a firm. There is some theoretical support for this risk mechanism, since a firm's ESG strategy may reduce the risk of reputation losses and/or potential litigation costs. The second mechanism relates changes in discount rates to investor preferences. If investors have preferences for environmental, social, and/or governmental issues (i.e., unrelated to standard financial preferences, such as risk and return) their demand will have an impact on a firm's stock price and, ultimately, on its discount rate. This is the main argument of Fama and French (2007). Analyzing and modeling non-financial tastes has attracted a rising attention in recent years. Pastor et al. (2020), Oehmke and Opp (2020) and Zerbib (2020) propose models in which agents have tastes for ESG stocks and analyze their implications on the discount rate. These classes of taste models imply a higher valuation and lower expected returns of ESG stocks. In sum, we address the ESG return puzzle by analyzing the cash-flow channel and the discount rate channel and distinguish the latter into a risk mechanism and a demand mechanism. This approach helps us to explain why realized returns differ from expected returns over periods during which the demand for ESG assets changes. Thereby, the decomposition may also help to forecast what returns for high versus low-ESG stocks can be expected in the future.

Empirically, we examine the two channels of the ESG return puzzle in the US stock market. Using a large sample of US companies from 2008 to 2018, we form a hedge portfolio which is long in *high* ESG stocks and short in *low* ESG stocks (hereafter: HL portfolio). Consistent with most previous empirical evidence (e.g., Kempf and Osthoff 2007; Li et al. 2019), the HL portfolio earns a positive realized return which equals about 2% p.a. We also find a negative expected return of about -0.5% p.a. for the HL portfolio which is which is about the same level reported by earlier studies such as, for example, Chava (2014) or El Ghoul et al. (2011, 2018).

We then analyze the cash-flow and discount rate channels of the unexpected return (i.e., difference between realized and expected return). We find no evidence that the cash-flow channel is able to explain a positive unexpected return of the HL portfolio. If at all, cash-flow news tends to make the unexpected return of the HL portfolio even larger. That is, cash-flow expectations of the HL portfolio are not revised upward. The discount rate channel, however, does provide an explanation for the ESG return puzzle. We find that discount rates of the HL portfolio have fallen over the sample period. We further show that the discount rate channel can be explained by investor demand but not by time varying risk attributes. To be specific, we find that an increasing share of investors with ESG preferences is positively correlated with a decreasing discount rate of the HL portfolio. This observation is compatible with the theoretical view of Fama and French (2007) that the demand of ESG investors drives prices of high-ESG stocks upward and, therefore, lowers their discount rates. In contrast, the risk mechanism does not receive empirical support, since we find that, on average, H stocks are less risky than L stocks are. In sum, the empirical results are compatible with the view that some investors have non-financial preferences linked to ESG. This view implies that expected future returns on the HL portfolio will be considerably lower than realized past returns if the demand for ESG stocks does not increase further. Several robustness checks provide evidence that the main conclusions are rather robust.

Our paper makes three main contributions. First, we apply the return decomposition approach of Campbell and Shiller (1988) to an ESG portfolio for the first time. Therefore, we are able to provide direct evidence of why the unexpected return of the HL portfolio is positive. Second, we provide evidence that the ESG return puzzle can mainly be described by discount rate news, while cash-flow news makes the puzzle—if at all—worse. Third, we link discount rate news of the HL portfolio to a risk channel and a demand channel. Our findings are compatible with the view that investor demand for ESG assets is strongly related to discount rate news. We therefore provide direct empirical support for the existence of a demand channel which has been suspected previously but not directly documented (e.g., Galema et al. 2008).

The paper is organized as follows. Section 2 provides a literature review of the expected and realized return of stocks in relation to ESG characteristics. In Section 3, we present the theoretical approach of how the ESG return puzzle is analyzed in this paper. Empirical issues are addressed in Section 4, and Section 5 presents the empirical results. Section 6 concludes the paper.

Literature review

The relation between stock returns and ESG characteristics has been extensively analyzed. We summarize the main findings documented in the literature that addresses the ESG return puzzle, i.e., the relation between ESG scores and stock returns (realized and expected).

ESG and realized returns

Empirical research on the US stock market supports the views that stocks with a good ESG rating tend to deliver a higher return than those with a bad rating. In a widely cited study, Kempf and Osthoff (2007) analyze a long/short strategy using KLD ratings (now MSCI) over the period from 1992 to 2004. They analyze different dimensions of ESG and combine them into an overall ESG score. A long/short portfolio which buys the best 5% (50%) of stocks and sells the worst 5% (50%) results in a risk-adjusted return of up to 8.70% (0.95%) per year. Their empirical results are close to ours reported below. The positive risk-adjusted return of ESG stocks also extends to other countries (e.g., Bauer et al. 2004; Brzeszczynski and McIntosh 2014). More recent studies include more detailed explanations for the positive abnormal investment return of ESG stocks. Eccles et al. (2014) argue that a potential explanation is that companies with a high sustainability score have better organizational processes in place than companies with low scores. Choi et al (2020) discuss that attention to specific ESG issue is a potential driver of returns. They find that carbon-intensive stocks tend to underperform when attention to climate change is high (as measured by Google search volume).

Although a larger number of studies support the main conclusion that ESG is positively related to realized returns of individual stocks,[1] there is also empirical evidence that a particular dimension of social irresponsibility, so-called sin stocks (e.g., alcohol, tobacco, and gaming), have earned a positive risk-adjusted return. Hong and Kacperczyk (2009) document that over the period 1965–2006, sin stocks outperformed comparable non-sin stocks by 26 basis points per month. The sin dimension seems to include environmental issues recently since the study of Bolton and Kacperczyk (2020) finds higher average stock returns for firms with higher carbon emissions.

ESG and expected returns

Expected returns of ESG stocks are more difficult to analyze, since they cannot be directly observed. To obtain estimates for the expected return, many studies use the forward looking concept of the implied cost of capital that was introduced by Claus and Thomas (2001) and Gebhardt et al. (2001). In the context of ESG, El Ghoul et al. (2011) analyze the implied cost of capital of a large sample of US firms from 1992 to 2007. Their study suggests that high-ESG companies (above median scores) have an expected return which is between 43 and 78 bps lower than that of low-ESG companies. In a recent study, El Ghoul et al. (2018) confirm these observations for a large sample of manufacturing firms from 30 countries. Focusing on the environmental profile of a company, Dhaliwal et al. (2020) report that the cost of equity tends to fall if a firm voluntarily reports on its social responsibility activities. Looking at bonds, ex ante cost of capital measures are easier to observe. Zerbib (2019), provides supporting evidence that the cost of debt is lower for green bonds relative to conventional bonds for example, although the difference is rather small (i.e., two basis points). However, these studies are not able to explain the ESG return puzzle, since they focus on the *level* of discount rates and not on the *changes* therein. From a theoretical point of view (see Merton 1973; Campbell and Vuolteenaho 2004), the consideration of discount rate *news* is, however, necessary to provide a complete picture of why the unexpected return of the HL portfolio is positive.

There are two different approaches to explain why good-ESG companies should have a lower cost of capital. The first approach argues that an ESG policy leads to a lower level of downside risk. For example, Hong and Kacperczyk (2009) argue that companies involved in "sin" businesses experience substantial litigation risk which should be reflected in a higher cost of capital. An alternative channel for explaining a lower downside risk is customer loyalty (e.g., Albuquerque et al. 2019). A higher loyalty may give companies more pricing power which finally reduces a firm's risk profile. Empirical support for such a downside risk channel is given, for example, by Ilhan et al. (2020). They find that carbon risk can be attributed to tail risk.

A second approach relates a company's cost of capital to investor preferences and tastes. A key assumption in the taste models is that agents derive utility in two forms. One reflects traditional financial utility (i.e., high return and low risk), the second is a non-financial (e.g., ethical) benefit. That is agents are happy to hold assets with positive ESG characteristics although they know that they have to sacrifice expected returns. Loosely speaking, investors make two trade-offs: risk versus return and a good conscience versus return. Riedl and Smeets (2017) provide empirical support for the latter trade-off. An early example of a taste model is Fama and French (2007). They analyze the effects of non-pecuniary tastes in a model with two investor types. The first investor has standard financial preferences, while the second investor has additional preferences for non-financial factors (such as tastes for ESG issues). Their model implies that a higher asset demand from the second type of investor can decrease the cost of capital for ESG firms. An alternative taste model is proposed by Pastor et al. (2020). In their model, the degree of how agents differ in their ESG preferences is key to explain a firm's cost of capital. In equilibrium, agents hold then a combination of three assets, the risk free asset, the market portfolio and an ESG hedge portfolio which is long in assets with positive ESG characteristics and short with negative characteristics. Accordingly,

[1] The issue whether the individual stock perspective extends to mutual funds, is discussed heterogeneously. Revelli and Viviani (2015) conduct a meta-analysis of 85 studies and 190 experiments and find that ESG does neither increase nor hurt portfolio performance. See also Liang and Renneboog (2020) for a more recent survey.

the expected return of an asset is determined by an asset's risk exposure to the market portfolio and the ESG hedge portfolio. Alternative models that consider investors with non-financial preferences have also been proposed by, for example, Luo and Balvers (2017), Oehmke and Opp (2020) and Zerbib (2020). In addition, multi-factor models which include an ESG factor have been suggested by, for example, Xiao et al. (2013) and Gregory et al. (2020).

Taste models can potentially explain the ESG return puzzle, if a changing demand from ESG investors is observed in the sample period. Then, realized returns can differ from their expectations. Thus, there is strong theoretical support for a negative (positive) expected (realized) return difference between high- and low-ESG stocks.

Theoretical approach

In the following, we provide the theoretical framework for how we investigate the ESG return puzzle.

Return decomposition of the HL portfolio

The efficient market hypothesis of Fama (1991) states that an asset's realized return should equal its expected return in the long-run. Over a short-term period, however, realized returns can differ from their expectations if unexpected news arrives at the market. The return decomposition framework of Campbell and Shiller (1988) formalizes this view. They show that a stock's unexpected return (UR_t) from $t-1$ to t, i.e., the difference between realized and expected return over one period, is

$$R_t - E_{t-1}(R_t) \equiv \mathrm{UR}_t = \mathrm{NCF}_t - \mathrm{NDR}_t, \tag{1}$$

where R_t is the log return from t to $t-1$, $E_{t-1}(R_t)$ is the expected log return at $t-1$, NCF_t and NDR_t are cash-flow news and discount rate news between $t-1$ an t, defined as

$$\mathrm{NCF}_t \equiv \Delta E_t \sum_{j=0}^{\infty} \rho^j \Delta D_{t+j} \text{ and } \mathrm{NDR}_t \equiv \Delta E_t \sum_{j=1}^{\infty} \rho^j R_{t+j}.$$

Thereby, D_t refers to the log dividend paid in t, and ρ is a number smaller than one resulting from the linearization approach. ΔE_t is the change in expectations from $t-1$ to t. In particular, we follow Campbell and Shiller (1988) and set $\rho^{12} = 0.96$; please note that our empirical analysis uses monthly data. Equation (1) states that a deviation of the realized return from its expectations can only be explained by changes in *expectations* of future cash flows and discount rates. In words, the value of a firm can only rise unexpectedly by either increasing cash-flow expectations and/or by applying lower discount rates to cash-flow expectations.

Equation (1) also holds for a long-short portfolio such as HL. Then, the difference of the unexpected return of the HL portfolio (highly ESG rated firms are denoted by superscript H and lowly ESG rated firms denoted by superscript L, specified in Section 4) is

$$\begin{aligned}
\underbrace{\mathrm{UR}_t^H - \mathrm{UR}_t^L}_{\equiv \mathrm{UR}_t^{\mathrm{HL}}} &= \underbrace{R_t^H - R_t^L}_{\equiv R_t^{\mathrm{HL}}} - \underbrace{\left(E_{t-1}(R_t^L) - E_{t-1}\left(R_t^H\right)\right)}_{\equiv E(R_t^{\mathrm{HL}})} \\
&= \underbrace{\left(\mathrm{NCF}_t^H - \mathrm{NCF}_t^L\right)}_{\equiv \mathrm{NCF}_t^{\mathrm{HL}}} - \underbrace{\left(\mathrm{NDR}_t^H - \mathrm{NDR}_t^L\right)}_{\equiv \mathrm{NDR}_t^{\mathrm{HL}}}. \qquad (2) \\
\mathrm{UR}_t^{\mathrm{HL}} &= \mathrm{NCF}_t^{\mathrm{HL}} - \mathrm{NDR}_t^{\mathrm{HL}}
\end{aligned}$$

Equation (2) is the framework for analyzing the ESG return puzzle (i.e., $\mathrm{UR}^{\mathrm{HL}} > 0$). According to Eq. (2), a positive $\mathrm{UR}^{\mathrm{HL}}$ implies either $\mathrm{NCF}^{\mathrm{HL}} > 0$ and $-\mathrm{NDR}^{\mathrm{HL}} > 0$ or $\mathrm{NCF}^{\mathrm{HL}} - \mathrm{NDR}^{\mathrm{HL}} > 0$. That is, two channels, a "cash-flow channel" and a "discount rate channel" provide a potential explanation for the ESG return puzzle. Considering the discount rate channel, we further distinguish two different mechanisms for why a stock's discount rate changes. The classical approach relates changes in the discount rate to changes in risk (e.g., Merton 1973). Alternatively, Fama and French (2007) show that if investor demand is driven by non-financial issues, discount rates may also vary with demand. Thus, we specify the discount rate channel with a risk mechanism and a demand mechanism (specified below). Thus, our framework analyzing the ESG return puzzle includes four steps:

1. Analysis of expected and realized returns of H and L stocks.
2. Formation of a hedge portfolio HL (long good ESF firms, short bad ESG-firms).
3. Decomposition of the difference between realized and expected return, $R_t^{\mathrm{HL}} - E_{t-1}(R_t^{\mathrm{HL}}) \equiv \mathrm{UR}_t^{\mathrm{HL}} = \mathrm{NCF}_t^{\mathrm{HL}} - \mathrm{NDR}_t^{\mathrm{HL}}$
4. Analysis of discount rate news ($-\mathrm{NDR}_t^{\mathrm{HL}}$) with risk channel and demand channel.

Cash-flow channel

A positive $\mathrm{UR}^{\mathrm{HL}}$ can be explained by positive cash-flow news of the HL portfolio. That is, high-ESG stocks should report *revisions* of future cash-flow expectations better than low-ESG stocks. We should emphasize that it is not the *level* of cash-flow expectations that determines the unexpected returns but changes in these expectations. The distinction between different levels and unexpected changes of cash flows is important, since previous research focused on levels of future cash-flow expectations. For example, Gregory et al. (2014) argue that growth prospects (i.e., levels) of high-ESG stocks are better than that of low-ESG stocks, and they interpret that those differences in growth expectations are the main driver of valuation differences between H and L stocks. However, differences in growth expectations and valuation differences do not imply that unexpected returns are higher for H than for L stocks and, therefore, do not help to explain the ESG return puzzle.

To give a simple intuitive example of the existence of the cash-flow channel: consider a situation where consumers change their preferences in favor of products of high-ESG-rated companies. If the perception of climate change leads consumers to prefer renewable energy (e.g., generated by wind and solar) to traditional energy (e.g., generated by coal or oil), changes in consumer preferences may lead to increasing demand for renewable energy products compared to traditional energy and potentially to higher prices and lower production costs. As a result, high-ESG companies are able to produce more and/or sell their products at higher margins compared to the situation before climate change was considered to be an important issue by consumers. In contrast, low-ESG companies sell less and their profit margins shrink. Accordingly, cash-flow expectations are revised upward for H companies and downward for L companies, since more ESG consumers switch to H firms' products.

The implications of this simple example are supported theoretically. The stakeholder theory of Jensen (2002) suggests that firms engage in a positive ESG policy to consider preferences of investors, employees, customers, and other stakeholders. Although these ESG activities may incur short-term costs, long-term benefits will outweigh them. Since cash-flow news in Eq. (1) includes both short-term and long-term cash-flow expectations, the net effect should be positive (i.e., $\mathrm{NCF}_t^{\mathrm{HL}} > 0$). A positive net effect can be further justified with a resource-based perspective of the company (e.g., Barney 1991). In such a context, McWilliam and Siegel (2011) argue that positive ESG activities are associated with a competitive advantage for the firm because they strengthen social relationships with stakeholders such as employees, customers, or suppliers. Once the relationship network has been built through ESG activities, it can become an irreplaceable strategic resource if its complexity is difficult to imitate (e.g., Colbert 2004). Then, if customers are willing to pay higher prices for products of a company with an ESG-friendly policy, the firm's cash flows may be higher than for companies which do not follow a positive ESG policy. However, there are also theories which predict the opposite, i.e., that a firm's orientation to ESG may lower its cash flows. For example, Friedman (1970) argues that ESG activities are mainly costs for a firm which are not necessarily compensated by increased revenues. If these costs unexpectedly increase over time, then analysts and investors may revise their cash-flow expectations downward for the HL portfolio (i.e., $\mathrm{NCF}_t^{\mathrm{HL}} < 0$). Thus, theoretical approaches may justify both positive and negative cash-flow news from the HL portfolios.

Empirically, there is some evidence that ESG and cash flows are actually related. For example, Godfrey et al. (2009) provide evidence that a company's ESG objectives result in good relationships with stakeholders, which in turn reduces a company's idiosyncratic risk profile and increases long-term cash-flow expectations. Similar arguments are put forward by other studies (e.g., Choi and Wang 2009; Gregory et al. 2014). Furthermore, research by Sen and Bhattachary (2001) provides evidence that consumers tend to pay more or to increase their purchase intention if they relate a company to good-ESG activities. Accordingly, Armstrong and Green (2013) argue from a stakeholder perspective that an ESG-friendly corporate policy is value enhancing, while the opposite is detrimental to a firm's value. Hong et al. (2016) provide empirical support for this view. However, most of the research mentioned here derives its conclusions from realized cash flows and does not consider cash-flow expectations. Thus, we interpret prior research as providing support for the existence of a cash-flow channel but without unambiguously deriving its direction. Since our own approach relies on expectations, we will provide direct evidence for the direction of the cash-flow channel.

Discount rate channel

Next to the cash-flow channel, the discount rate channel may potentially explain the ESG return puzzle. A positive unexpected return of the HL portfolio, then, requires positive discount rate news. That is, discount rates of H stocks should fall relative to L stocks. As argued above, discount rates can vary by reason of changing risk characteristics or changing investor demand.

Risk mechanism

To justify time-varying discount rates requires an intertemporal asset pricing approach. We use the intertemporal capital asset pricing model (ICAPM) of Merton (1973). On the basis of the ICAPM, Campbell and Vuolteenaho (2004) have specified a two-beta version of the traditional capital asset pricing (CAPM) beta (Sharpe 1964; Lintner 1965) which decomposes the CAPM-beta into a cash-flow beta and a discount rate beta:

$$\beta_t^{\mathrm{CAPM}} = \beta_t^{\mathrm{CF}} + \beta_t^{\mathrm{DR}}, \tag{3}$$

where

$$\beta_t^{\mathrm{CAPM}} \equiv \mathrm{cov}_t\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, \mathrm{UR}_{t+1}^{M}\right) / \mathrm{var}_t\left(\mathrm{UR}_{t+1}^{M}\right), \quad \text{CAPM beta,}$$

$$\beta_t^{\mathrm{CF}} \equiv \mathrm{cov}_t\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, \mathrm{NCF}_{t+1}^{M}\right) / \mathrm{var}_t\left(\mathrm{UR}_{t+1}^{M}\right), \quad \text{Cash - flow beta,}$$

$$\beta_t^{\mathrm{DR}} \equiv \mathrm{cov}_t\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, -\mathrm{NDR}_{t+1}^{M}\right) / \mathrm{var}_t\left(\mathrm{UR}_{t+1}^{M}\right), \quad \text{Discount - rate beta.}$$

Then, the ICAPM implies for the expected return of the HL portfolio

$$E_t\left(R_{t+1}^{\mathrm{HL}}\right) - R_{f,t} + \frac{\sigma_{\mathrm{HL},t}^2}{2} = \gamma \cdot \beta_t^{\mathrm{CF}} \cdot \sigma_{M,t}^2 + \beta_t^{\mathrm{DR}} \cdot \sigma_{M,t}^2, \tag{4}$$

where γ is the coefficient of relative risk aversion and $\sigma_{M,t}^2$ is the conditional variance of the market portfolio. The factor $\frac{\sigma_{\mathrm{HL},t}^2}{2}$ on the left hand side is one-half of the variance of the return of the HL portfolio in order to adjust for Jensen's inequality. If the ICAPM holds, the ESG risk premium puzzle implies that a higher return can be earned if the HL portfolio's cash-flow beta β_t^{CF} and/or discount rate beta β_t^{DR} is larger than zero.

Betas in Eq. (3) are conditional on time t. Therefore, empirical implementation of the ICAPM requires a specification of how discount rates vary overtime. We follow a simple approach, as in Botshekan et al. (2012), and assume that betas are different in up and down markets. To be specific, we model the discount rate beta as $\beta_t^{\mathrm{DR}} \equiv \beta^{\mathrm{DR}} + \beta^{\mathrm{DR+}}$, where $\beta^{\mathrm{DR+}}$ is the additional up beta when the unexpected return of the market portfolio is positive (i.e., $\mathrm{UR}_t^M > 0$). Thus, discount rate betas vary through time, depending on the unexpected return of the market portfolio and equal β^{DR} in down markets and $\beta^{\mathrm{DR}} + \beta^{\mathrm{DR+}}$ in up markets. Similarly, we model the cash-flow beta and obtain

$$\begin{aligned} \beta^{\mathrm{CF+}} &= \mathrm{cov}\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, \mathrm{NCF}_{t+1}^{M} | \mathrm{UR}_{t+1}^{M} > 0\right) / \mathrm{var}\left(\mathrm{UR}_{t+1}^{M}\right) \quad \text{additional up cash - flow beta,} \\ \beta^{\mathrm{DR+}} &= \mathrm{cov}\left(UR_{t+1}^{HL}, -NDR_{t+1}^{M} | UR_{t+1}^{M} > 0\right) / \mathrm{var}\left(UR_{t+1}^{M}\right) \quad \text{additional up discount rate beta.} \end{aligned} \tag{5}$$

Demand mechanism

Besides risk, demand is an alternative explanation for why discount rates may change. If, for example, preferences of investors change over time for reasons other than risk and return (e.g., ESG preferences increase), this may drive up prices and, accordingly, discount rates will fall. Fama and French (2007) present a general analysis of how investor demand affects asset prices and their expected returns. The main channel is that investors derive utility from holding specific assets other than return. For example, by holding an ESG asset, an investor may feel that she is doing some good, which in turn increases her utility. If, over time, the taste for ESG increases, an increasing demand for H assets drives their prices up relative to L stocks ($R_t^{\mathrm{HL}} > 0$) and, simultaneously, drives their expected returns down ($-\mathrm{NDR}_t^{\mathrm{HL}} > 0$), and vice versa. Similar taste channels have been analyzed by Gregory et al. (2020), Oehmke and Opp (2020), Pastor et al. (2020) and Zerbib (2020). In general, taste models imply that the expected return of the HL portfolio is a function of investor demand, i.e.,

$$E\left(R_t^{\mathrm{HL}}\right) = f\left(\mathrm{demand}_t\right), \tag{6}$$

where demand is the percentage of ESG investors to all investors. This mechanism works even in the case of constant cash-flow expectations. Assuming for simplicity that f in Eq. (6) is linear implies that

$$-\mathrm{NDR}_t^{\mathrm{HL}} = \alpha + \beta^{\mathrm{demand}} \cdot \Delta\mathrm{demand}_t + \varepsilon_t. \tag{7}$$

If investor demand for ESG assets drives their prices, the slope parameter β^{demand} should be larger than zero. Such a demand channel has not been analyzed empirically in previous studies. In a different context, however, demand effects of ESG investors have been addressed. For example, Robinson et al. (2011) report that stocks of firms added to the Dow Jones Sustainability World Index (DJSI) experience a sustained increase in their share price, while stocks deleted from the index had a temporary decrease in their valuation. They explain these price effects by demand from investors with ESG preferences. However, their event study approach focuses on the demand for individual stocks and not ESG stocks in general. Also, they are not able to differentiate between cash-flow news and discount rate news, which, theoretically, are important components of the unexpected return.

Empirical approach

How we measure cash-flow news and discount rate news

We apply the return composition of Campbell and Shiller (1988) in order to explain the past return of ESG and non-ESG stocks. The question arises of how to operationalize the decomposition in Eq. (1). The literature follows two approaches: the vector autoregressive (VAR) approach proposed by Campbell and Shiller (1988) and the use of a valuation model employing analysts' earnings forecasts (e.g., Chen et al. 2013). The VAR-based approach is usually applied at the aggregate market level (and not at the individual company level as in this study). Further, it has been criticized, for example, by Chen and Zhao (2009), since its resulting news estimates are unstable and heavily dependent on the state variables included in the predictive VAR model.

Chen et al. (2013) also provide empirical evidence that the approach using a valuation model and earnings forecasts is preferable for identifying the underlying driving forces of the unexpected return at both the firm and the aggregate market level. We therefore use the approach based on analysts' forecasts and compute the implied cost of capital (ICC) as an estimate of the expected return, which is also widely used in the ESG literature (e.g., El Ghoul et al. 2011, 2018; Chava 2014). We estimate the ICC by using various models, i.e., Gebhardt et al. (2001), Claus and Thomas (2001), Easton (2004), Ohlson and Juettner-Nauroth (2005), and the average cost of capital estimates from those four models. Tang et al. (2014) argue that the model of Gebhardt et al. (2001), hereafter GLS, seems to have superior characteristics to the alternative models. Therefore, we consider GLS as our base case model, which we describe in more detail below and explain how it is used to calculate discount rate news. For a description of the alternative models, we refer the reader to the original papers or to the short descriptions given in El Ghoul et al. (2011).

The GLS model assumes clean surplus accounting and expresses the current share price in terms of expected returns on equity $E_t(\mathrm{ROE}_{t+j})$ and book values B_{t+j} for fiscal year j ahead of t. The infinite forecast horizon is divided into three time periods: an explicit forecast period for the ROE for the next three fiscal years, a convergence period between fiscal year four and twelve during which the ROE is expected to converge to the median industry ROE, and a period after year twelve in which the expected ROE is assumed to be constant. Similar assumptions are made for the dividend payout ratio. Then, the current stock price P is

$$P_t = B_t + \sum_{j=1}^{11} \frac{E_t(\mathrm{ROE}_{t+j}) - \mathrm{ICC}_{t+j}^{\mathrm{GLS}}}{(1+\mathrm{ICC}_t^{\mathrm{GLS}})^j} \cdot B_{t+j-1} + \frac{E_t(\mathrm{ROE}_{t+12}) - \mathrm{ICC}_t^{\mathrm{GLS}}}{\mathrm{ICC}_t^{\mathrm{GLS}} \cdot (1+\mathrm{ICC}_t^{\mathrm{GLS}})^{11}} \cdot B_{t+11}. \tag{8}$$

Using the implied cost of capital factor $\mathrm{ICC}_t^{\mathrm{GLS}}$ at time t from Eq. (8), the discount rate news factor is

$$\mathrm{NDR}_t^{\mathrm{GLS}} \equiv \Delta E_t \sum_{j=1}^{\infty} \rho^j \mathrm{ICC}_t^{\mathrm{GLS}} = \frac{\Delta \mathrm{ICC}_t^{\mathrm{GLS}}}{1-\rho}. \tag{9}$$

Using the GLS expected return in $t-1$, $\mathrm{ICC}_{t-1}^{\mathrm{GLS}}$ as an estimator for $E_{t-1}(R_t)$ in Eq. (1), cash-flow news can then be backed out easily.

Data

To implement our approach, we use various data sources. The data source for firm-level ESG scores is MSCI. MSCI has acquired KLD Research & Analytics, which has been in many studies the main database (e.g., Kempf and Osthoff 2007; El Ghoul et al. 2011, 2018). Thereby, MSCI and KLD have been one of the first suppliers of ESG ratings next to Vigeo-Eiris which has been acquired by Moody's recently (see Berg et al. 2019). From an investor perspective, MSCI, next to Sustainalytics, is the most favored ESG rating provider (see SustainAbility 2020). In a recent Extel survey, MSCI has been voted to be number one in ESG research (Extel Survey 2019).

The use of one rating provider can be criticized since the number of ESG rating agencies has increased substantially over the last years. According to Li and Polychronopoulos (2020), there are currently at least as 70 different sources for ESG ratings. Berg et al. (2019) among others provide evidence that ESG ratings for an individual company can diverge considerably even among top rating agencies. The main reason for this divergence is that there is not yet common accepted method how to measure ESG. However, this divergence problem seems to be less severe in the top minus bottom portfolio approach (using quantile ranks) as shown by Berg et al. (2019). They report that the implied correlation across different ESG ratings is about 80% by using a quantile rank count approach. The focus on MSCI ESG ratings can therefore be justified by practical end empirical arguments. However, one should consider this limitation when interpreting empirical results.

ESG scores of MSCI can be broken down into individual environmental (E), social (S), and governance (G) scores. While the ESG score is industry-adjusted (each industry has a median score of 5, the score ranges between 0 and 10), the individual E, S, and G scores are measured on an absolute scale (also ranging between 0 and 10). We primarily use the industry-adjusted ESG score and present results from using individual E, S, and G scores in the robustness section. We estimate the ICC using four different models which requires the use of earnings estimates (e.g., from financial analysts). We use the mean estimates provided by I/B/E/S and follow Gebhardt et al. (2001), who require that a company has 1- and 2-year-ahead consensus earnings estimates and a positive book value. We obtain earnings estimates and stock data (total return, price and book value) from Datastream.[2] Monthly data are collected at the end of each month and yearly data at the end of each calendar year. We use all US stocks for which data are available and delete small stocks that have a market capitalization of less than 0.1% of the median market capitalization of all stocks.

We calculate return data for individual stocks included in the HL portfolio. Therefore, at the end of each month we rank all stocks on their latest ESG score. Based on this

[2] We thank Quoniam Investment for providing the data and the capacity to do the calculations.

Table 1 ESG return puzzle, cash-flow channel and discount rate channel (equal weighted portfolios 2008–2018)

	GLS (1) (%)	CT (2) (%)	OJ (3) (%)	E (4) (%)	Average (5) (%)
Panel A: Average returns of portfolios H and L					
R^{H}	6.68	6.68	7.17	6.64	7.08
R^{L}	4.42	4.37	5.31	5.02	5.10
Panel B: Average returns and risk-adjusted returns (alphas) of portfolio HL					
R^{HL}	2.26	2.31	1.86	1.62	1.98
$R^{\mathrm{HL}}_{\mathrm{FF3}}$	3.05	3.11	2.61	2.28	2.76
$R^{\mathrm{HL}}_{\mathrm{C4}}$	3.06	3.11	2.61	2.29	2.76
$R^{\mathrm{HL}}_{\mathrm{FF5}}$	2.64	2.67	2.22	1.89	2.34
$R^{\mathrm{HL}}_{\mathrm{FF6}}$	2.63	2.66	2.21	1.89	2.32
Panel C: Decomposed average returns of portfolio HL					
$\mathrm{ER}^{\mathrm{HL}}$	−0.42	−0.14	−0.35	−0.60	−0.36
$\mathrm{UR}^{\mathrm{HL}}$	2.68	2.45	2.21	2.23	2.34
$\mathrm{NCF}^{\mathrm{HL}}$	0.25	−1.14	−1.25	−0.09	0.23
$-\mathrm{NDR}^{\mathrm{HL}}$	2.43	3.59	3.46	2.31	2.11

This table reports average returns of a portfolio with above or equal to median ESG scores (portfolio H) and a portfolio with below median ESG scores. Returns of each in the portfolios H and L are equally weighted; the portfolio HL is long in H and short in L. Expected returns (ER) are derived from the four different models developed by Gebhardt et al. (2001), Claus and Thomas (2001), Ohlson and Juettner-Nauroth (2005), and Easton (2004), respectively. "Average" refers to the mean expected return across the four models. UR refers to the unexpected return which is decomposed into cash-flow news (NCF) and discount rate news (−NDR) using the method of Campbell and Shiller (1988). All returns are annualized using monthly returns in the calculations

score, we form equal-weighted portfolios[3] and hold these portfolios unchanged until the end of the following month. The high-rated portfolio, denoted by superscript H, consists of the top 50% of all stocks with the highest ESG ratings; the low-rated portfolio is denoted by superscript L and consists of the bottom 50% with the lowest ESG ratings of all stocks. Then, the superscript HL denotes the return of the long-short portfolios in period t, i.e., R_t^{HL}. The average return over time is denoted by $\overline{R^{\mathrm{HL}}}$. The time period covered for our empirical analysis is January 2008 to July 2018. Factor returns used in the risk-adjusted regressions are obtained from Kenneth French's website.

Results

Empirical return decomposition

Table 1 summarizes average returns of the H and L portfolio (Panel A), average returns and alphas of the HL portfolio (Panel B), and results of the return decomposition in Panel C. Panel A shows that H stocks tend to achieve a higher return than L stocks, $\overline{R^{\mathrm{H}}} > \overline{R^{\mathrm{L}}}$. Looking at the first column (expected returns are derived from the GLS model), an equal-weighted portfolio of H-stocks delivers an average realized of 6.68% p.a., while the corresponding portfolio of L-stocks earns just 4.42% p.a. Panel B displays the difference between the two portfolios, $\overline{R^{\mathrm{HL}}}$, which is 2.26% p.a. for the GLS approach. Using alternative approaches for modeling expected returns (columns (2) to (5)) provides similar results; the average return of the HL portfolio varies between 1.62 and 2.31% across all ICC models. We also adjust the average return of the HL portfolio with common multi-factor models such as Fama and French (1993), denoted by FF3, Carhart (1997), C4, Fama and French (2015), FF5, and Fama and French (2018), FF6. These models consider—next to the market beta—various risk factors, such as size, value, momentum, investment, and profitability. In general, the average HL return is only marginally affected by those risk-adjustments. For example, using the base case model GLS, the average HL return is 2.26% p.a., while the FF6 risk-adjusted HL return is even higher (2.63% p.a.). Thus, the average HL returns cannot be attributed to known risk factors and the size of risk-adjusted average returns is similar to recent studies (e.g., Kempf and Osthoff 2007; Li et al. 2019). These observations provide confirming empirical support for the first part of the ESG return puzzle ($\overline{R^{\mathrm{HL}}} > 0$).

[3] We use equally weighted portfolios in our base case scenario. However, all conclusions presented in the empirical section do hold for a market capitalization weighted portfolio (see also the robustness section).

In Panel C, we summarize the results of the return decomposition. The average expected return of the HL portfolio, denoted by $E\left(R^{\mathrm{HL}}\right)$, delivers a negative value in the range between -0.60 and -0.14% across the various specifications of the expected return model. Thereby, the second part of the ESG return puzzle is also observed ($E\left(R^{\mathrm{HL}}\right) > 0$), which implies that good-ESG companies tend to have a lower expected return (i.e., cost of capital) than bad ESG firms do. This observation is also compatible with earlier studies. For example, El Ghoul et al. (2011) find an average difference in cost of capital estimates between above and below median ESG companies of between -0.78 and -0.31% over the period from 1992 to 2007. Thus far, the summary statistics show that in our sample a positive $\overline{R^{\mathrm{HL}}}$ and a negative $E\left(R^{\mathrm{HL}}\right)$ are observed, which leads to an average unexpected returns of the HL portfolio between 2.21 and 2.68% p.a. We conclude that the sensitivity of specifying the estimation approach of the expected stock return seems to have a minor impact on the existence of the ESG return puzzle.

Estimates of cash-flow news and discount rate news suggest that positive unexpected returns of the HL portfolio can primarily be attributed to the discount rate channel. Discount rate news varies between 2.11 and 3.59% and explains to a large extent the unexpected return of the HL portfolio. The cash-flow channel, however, cannot explain the HL portfolio's positive unexpected return. Over the sample period cash flow news varies around zero, ranging from -1.14 to $+0.25\%$ (depending on which cost of capital model is applied). Thus, it seems unlikely that cash-flow news is a major driver of the ESG return puzzle, whereas discount rate news seems to be the main explanation of the puzzle. Discount rates (i.e., expected returns) of good-ESG companies have fallen to a larger extent than those of bad ESG companies. In the following section, we analyze this discount rate channel in more detail.

We also calculate the summary statistics for a value-weighted portfolio and a more recent period from 2013 to 2018 to consider the observation of Li at al. (2019), since they report that alphas of ESG portfolios have fallen in recent years. However, we do not observe this pattern and the conclusions from using a value-weighted HL portfolio and a more recent sample are the same as in Table 1. Therefore, we present details of these statistics in a robustness analysis.

Discount rate channel

The last section has shown that the unexpected return of the HL portfolio is primarily driven by discount rate news. In the following section, we analyze the two potential mechanisms of the discount rate channel, namely the risk mechanism and the demand mechanism.

Table 2 Cash-flow betas and discount rate betas of the portfolio HL

	GLS (1)	CT (2)	OJ (3)	E (4)	Average (5)
Panel A: Risk mechanism					
β^{CF}	0.000	-0.037	-0.049	-0.064	-0.047
	(0.001)	(-2.483)	(-5.266)	(-3.909)	(-1.819)
$\beta^{\mathrm{CF+}}$	-0.106	-0.003	-0.135	-0.197	-0.158
	(-1.006)	(-1.168)	(-1.874)	(-2.252)	(-0.447)
β^{DR}	-0.064	-0.029	-0.103	-0.107	-0.073
	(-4.441)	(-2.739)	(-3.733)	(-16.313)	(-6.659)
$\beta^{\mathrm{DR+}}$	-0.106	-0.003	-0.135	-0.197	-0.158
	(-1.920)	(-1.088)	(-5.499)	(-5.081)	(-2.869)
Panel B: Demand mechanism					
β^{demand}	26.744	39.912	59.519	17.889	28.321
	(2.769)	(2.664)	(2.585)	(1.816)	(2.250)

This table reports decomposed beta factors of the portfolio HL, which is long in H stocks (stocks with above or equal to median ESG scores) and short in L stocks with below median ESG scores). Cash-flow and discount rate betas are computed following Campbell et al. (2010) as $\beta_t^{\mathrm{CF}} \equiv \mathrm{cov}_t\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, \mathrm{NCF}_{t+1}^{M}\right)/\mathrm{var}_t\left(\mathrm{UR}_{t+1}^{M}\right)$ and $\beta_t^{\mathrm{DR}} \equiv \mathrm{cov}_t\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, -\mathrm{NDR}_{t+1}^{M}\right)/\mathrm{var}_t\left(\mathrm{UR}_{t+1}^{M}\right)$, while the computation of the additional up betas follows Botshekan et al. (2012) as $\beta^{\mathrm{CF+}} = \mathrm{cov}\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, \mathrm{NCF}_{t+1}^{M} | \mathrm{UR}_{t+1}^{M} > 0\right)/\mathrm{var}\left(\mathrm{UR}_{t+1}^{M}\right)$ and $\beta^{\mathrm{DR+}} = \mathrm{cov}\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, -\mathrm{NDR}_{t+1}^{M} | \mathrm{UR}_{t+1}^{M} > 0\right)/\mathrm{var}\left(\mathrm{UR}_{t+1}^{M}\right)$

Unexpected returns (UR) cash-flow news (NCF) and discount rate news ($-$NDR) are derived by using the method of Campbell and Shiller (1988). Thereby, expected returns (ER) are from the four different models developed by Gebhardt et al. (2001), Claus and Thomas (2001), Ohlson and Juettner-Nauroth (2005), and Easton (2004), respectively. "Average" refers to the mean expected return across the four models. The demand beta β^{demand} is obtained from the regression $-\mathrm{NDR}_t^{\mathrm{HL}} = \alpha + \beta^{\mathrm{demand}} \cdot \Delta\mathrm{demand}_t + \varepsilon_t$ using quarterly returns over the sample period 2008 to 2018. t-values in parentheses

Risk mechanism

The risk mechanism of the discount rate channel implies that the HL portfolio should have a positive cash-flow beta and/or a positive discount rate beta. That is, H stocks should display a larger beta risk than L stocks do. A positive beta for the HL portfolio is necessary in our sample period, since the market portfolio delivered a positive unexpected return of about 1% p.a., mainly because discount rates of the market portfolio have fallen. Panel A of Table 2 provides no empirical support for such a risk mechanism. Looking at our base case model GLS in column (1), neither part of the CAPM-beta (cash-flow beta or discount rate beta) of the HL portfolio is larger than zero. Also, additional up-betas ($\beta^{\mathrm{CF+}}$ and $\beta^{\mathrm{DR+}}$) are not significantly greater than zero. For example, using the GLS model, the estimates of the cash-flow betas, β^{CF} and $\beta^{\mathrm{CF+}}$, are close to zero (0.000 and -0.106). Discount rate betas are estimated to be even smaller, and their point estimates ($\beta^{\mathrm{DR}} = -0.06$ and $\beta^{\mathrm{DR+}} = -0.106$) suggest that H stocks are *not riskier* than L stocks are. Rather, the significant negative estimates of discount rate betas imply that H stocks are *less risky* than L stocks are. The alternative ICC models produce similar estimates and in particular the two parts of the discount rate beta are more than two standard deviations below zero for almost all specifications. To put the estimate of the discount rate beta into an economic perspective, discount rate news of the market portfolio of 1% translates into an unexpected return of the HL portfolio of about $-0.17\% = (-0.064 – 0.106) \cdot 1\%$.

In sum, results presented in Table 2 are compatible with the view that high-ESG stocks are less risky than low-ESG stocks are. Although this evidence provides no explanation for the ESG return puzzle, it supports many studies from the management literature which argue that following ESG-strategies is less risky than not following them (e.g., because of reputation risk).

Demand mechanism

The demand mechanism of the discount rate channel implies that an increasing demand of ESG investors should lower the expected return of the HL portfolio (ceteris paribus). Further, investor demand implies a positive relation between discount rate news of the HL portfolio and changes in the demand of ESG investors, resulting in a positive demand beta. We approximate the demand of ESG investors by using assets under management (*AuM*) invested ESG equity funds (in the USA) and assets under management in all US equity funds (*TOT*),[4] i.e., $\text{demand}_t \equiv \text{AuM}_t(\text{ESG}) \big/ \text{AuM}_t(\text{TOT})$. Then, we run regression (7) and summarize the estimates of the demand beta in Panel B of Table 2.

Consistent with the prediction of demand models like that of Fama and French (2007), the demand beta β^{demand} is estimated to be consistently larger than zero. For example, using the GLS model to estimate expected returns (see column (1)), the slope estimate equals 26.744 (t-value of 2.769). An increasing share of ESG investors is therefore associated with a fall in discount rates of the HL portfolio (which in turn increases the stock price of H stocks relative to L stocks). Also, changes in the ESG-ratio can explain about 15% of the quarterly variation in the discount rate news of the HL portfolio. To put the slope estimate of 26.744 into an economic perspective, an increase of our demand proxy by 10%-points is accompanied by discount rate news of 2.674%. Although the point estimate of the demand beta varies across the different expected return models (see columns (2) to (5)), all estimates are larger than zero and in most cases the estimate is more than two standard errors above zero. In sum, changes in the share of ESG investors can explain a substantial variation in discount rate news of the HL portfolio, which is predicted by the Fama and French (2007) model.

Robustness of results

In this chapter, we present evidence that the results of the previous section are robust to changes in various assumptions in our base case approach. Thereby, we consider non-monotonic relations between ESG scores and discount rate news, we analyze alternative cutoffs when forming the HL portfolio, we investigate individually the E, S, and G scores, and we use value-weighted HL portfolios and alternative time periods. We display the results of using the GLS model for estimating discount rate news, since Tang et al. (2014) argue that the ICC approach of Gebhardt et al. (2001) seems to be the best proxy for a stock's expected return. However, the choice of this model is rather insensitive to the main conclusions about the robustness of the results that we present in this section.

Non-monotonic relationship between ESG scores and discount rate news

The demand channel implies a monotonic relationship between ESG and discount rates. However, there may be alternative relationships between ESG and return. Kim and Statman (2012) argue that the relationship between a firm's ESG policy and its stock return is nonlinear and they propose the existence of an optimal level of ESG. Then, adjustments to the optimal level should be compensated by a higher stock return. For example, if a company has overinvested in ESG, reducing the ESG investments may increase the value of the company. Using Kim and Statman's (2012) reasoning implies that both a fall (from a high level) and an increase (from a low level) in ESG scores can be associated with positive unexpected returns if a company adjusts to the optimal level of ESG. This reasoning suggests an inverted U-shape relation between changes in ESG and unexpected stock returns.[5] In contrast, the demand channel implies that only positive changes are valued and, in particular, those of high-ESG companies.

Therefore, we consider a potential adjustment channel by limiting our sample to those companies for which a change in their ESG score is observed. We form a long/short portfolio of stocks based on changes in ESG scores, denoted by ΔHL. Thereby, ΔH refers to stocks with a positive change,

[4] We obtain AuM figures from Bloomberg on a quarterly basis. Accordingly, we adjust discount rate news in the HL spread to quarterly data.

[5] The question about the optimal level of ESG is, however, discussed controversially. While Kim and Statman (2012) assume that a medium level of ESG is optimal, Barnett (2007) argues that only a real commitment (with potentially higher costs) to ESG is valued by customers. This implies that only high levels of ESG are ultimately valued, while medium investments in ESG or a reduction of ESG investments do not pay off.

ΔL refers to stocks with a negative change in their ESG scores, and ΔHL is the long/short portfolio. The question now arises of over which time horizon those changes in ESG scores should be measured. Empirically, Gregory and Whittaker (2013) observe that scores measuring the ESG dimension are relatively stable through time and that month-to-month changes in ESG scores are rather rare. Accordingly, using monthly changes would result in long and short portfolios, which consist of just a few (or in some months even of zero) stocks in the ΔHL portfolio. We therefore consider changes in ESG scores over one year as a compromise between a timely measure of ESG changes and an appropriate number of stocks in the ΔHL portfolio.

Table 3 presents the results of this exercise in column (1). The ΔHL portfolio delivers a positive unexpected return of 2.55% p.a., which suggests that positive changes in ESG scores are valued higher than negative changes (see Panel A). The unexpected return can thereby be primarily explained by discount rate news (which equals 2.69% p.a.). Thus, a long/short portfolio of stocks based on *changes* in ESG scores has similar return characteristics to an HL portfolio which is derived from the *level* of ESG scores. Also, systematic risk factors (displayed in Panel B) are not able to explain the ESG return puzzle. Most estimates of the decomposed beta factors are negative. In many cases, the associated *t*-values suggest a significant negative relationship between the ΔHL portfolio's discount rate news and that of the market portfolio. However, the positive demand beta of the portfolio ΔHL is compatible with the demand mechanism (see Panel C). The demand beta is about 53 and more than two standard errors above zero. These observations support the previous conclusions that investors' taste for ESG investments (demand mechanism) is a potential explanation for ESG stock returns.

Next, we consider the adjustment channel of the Kim and Statman (2012) by looking at ESG changes below/above the median ESG score of all companies. Therefore, we further split the ΔHL sample into two groups. The first group displays an ESG score above the median across all firms (see column (2)); the second group has an ESG score below the median score (column (3)). We assume that the median ESG score is an appropriate measure for the optimal level of ESG.[6] Within each of the two groups, we form a long/short portfolio according to ΔHL and denote it by ΔHL^{above} for the first group and ΔHL^{below} for the second group. Following the adjustment channel of Kim and Statman (2012), a negative relation between stock returns and changes in ESG should be observed in the first group ($UR^{\Delta HL^{above}} < 0$) and a positive relation in the second group ($UR^{\Delta HL^{below}} > 0$), because companies with below median ESG scores should invest in ESG to create value, while companies with above median ESG scores should reduce investments in ESG, assuming that a median ESG score is optimal. Results in columns (2) and (3) are in contrast to the adjustment channel of Kim and Statman (2012). In particular, the ΔHL^{below} delivers a negative unexpected return of −0.33% per annum, while the ΔHL^{above} earns an

Table 3 The ESG return puzzle, cash-flow betas, and discount rate betas of the portfolio ΔHL (i.e., using changes in ESG scores)

	ΔHL(1)	ΔHL^{above}(2)	ΔHL^{below}(3)
Panel A: ESG return puzzle: cash-flow channel and discount rate channel			
R^{HL}	2.55%	6.13%	−0.76%
ER^{HL}	−0.14%	−0.12%	−0.43%
UR^{HL}	2.69%	6.26%	−0.33%
NCF^{HL}	−4.46%	−4.30%	−2.84%
$-NDR^{HL}$	7.15%	10.56%	2.51%
Panel B: Risk mechanism			
β^{CF}	0.016	0.069	-0.011
	(0.311)	(0.544)	(0.090)
β^{CF+}	−0.101	−0.778	0.142
	(−0.180)	(−0.277)	(0.060)
β^{DR}	−0.072	−0.581	0.111
	(−2.962)	(−16.754)	(1.918)
β^{DR+}	−0.101	−0.778	0.142
	(−1.398)	(−4.523)	(1.051)
Panel C: Demand mechanism			
β^{demand}	52.762	77.231	11.422
	(2.189)	(1.787)	(0.231)

This table reports average returns of the portfolio HL, which is long in H stocks (stocks with above or equal to median ESG scores) and short in L stocks with below median ESG scores). Returns of each in the portfolios H and L are equally weighted; the portfolio HL is long in H and short in L. Expected returns (ER) are derived from the model developed by Gebhardt et al. (2001). "Average" refers to the mean expected return across the four models. UR refers to the unexpected return, which is decomposed into cash-flow news (NCF) and discount rate news (−NDR) using the method of Campbell and Shiller (1988). All returns are annualized using monthly returns in the calculations. Cash-flow and discount rate betas are computed following Campbell et al. (2010) as $\beta_t^{CF} \equiv \mathrm{cov}_t\left(UR_{t+1}^{HL}, NCF_{t+1}^{M}\right)/\mathrm{var}_t\left(UR_{t+1}^{M}\right)$ and $\beta_t^{DR} \equiv \mathrm{cov}_t\left(UR_{t+1}^{HL}, -NDR_{t+1}^{M}\right)/\mathrm{var}_t\left(UR_{t+1}^{M}\right)$, while the computation of the additional up betas follows Botshekan et al. (2012) as $\beta^{CF+} = \mathrm{cov}\left(UR_{t+1}^{HL}, NCF_{t+1}^{M} | UR_{t+1}^{M} > 0\right)/\mathrm{var}\left(UR_{t+1}^{M}\right)$ and $\beta^{DR+} = \mathrm{cov}\left(UR_{t+1}^{HL}, -NDR_{t+1}^{M} | UR_{t+1}^{M} > 0\right)/\mathrm{var}\left(UR_{t+1}^{M}\right)$

The demand beta β^{demand} is obtained from the regression $-NDR_t^{HL} = \alpha + \beta^{demand} \cdot \Delta demand_t + \varepsilon_t$ using quarterly returns over the sample period 2008 to 2018. *t*-values in parentheses

[6] If we limit the first group to companies with an ESG score above the 75%-quantile and the second group to companies with ESG scores below the 25%-quantile, we obtain qualitatively similar results.

unexpected return of 6.26%. A large part of this return can be attributed to the discount rate news factor. Furthermore, the discount rate factor cannot be explained by traditional risk factors, since discount rate betas are negative for ΔHL and $\Delta\text{HL}^{\text{above}}$ portfolios and positive for the $\Delta\text{HL}^{\text{below}}$ portfolio (see Panel B). Notice that all estimates of cash-flow betas are not distinguishable from zero. For example, the discount rate beta for the ΔHL portfolio is -0.072 and for the $\Delta\text{HL}^{\text{above}}$ portfolio it is -0.581, both significantly smaller than zero. Thus, stocks with positive changes in ESG scores above the median score can be characterized as less risky (with respect to market discount rate news) than stocks with positive changes, although they deliver a higher unexpected return. In contrast, the estimate of the discount rate beta for the $\Delta\text{HL}^{\text{below}}$ portfolio is 0.111 and almost two standard errors above zero. This estimate implies that stocks with an ESG score below the median which improve their scores are considered to be riskier than those which worsen their score. However, this risk is compensated by a negative unexpected return over the sample period.

Panel C strengthens the view that the coefficients are compatible with the demand channel. Using the model of Fama and French (2007) implies a larger demand beta for the $\Delta\text{HL}^{\text{above}}$ portfolio than for the $\Delta\text{HL}^{\text{below}}$ portfolio, since next to the level of ESG also changes in ESG should drive the investor's asset demand. The estimates of the regression coefficients are compatible with this implication. The estimate of demand beta for the $\Delta\text{HL}^{\text{above}}$ portfolio is about 77 (and marginally significant), while it is just 11 (not significant) for the $\Delta\text{HL}^{\text{below}}$ portfolio. In addition, the demand regression displays a larger adjusted R^2 for the $\Delta\text{HL}^{\text{above}}$ portfolio than for the $\Delta\text{HL}^{\text{below}}$ portfolio (7.75% versus 1.65%, not shown in Table 3). Thus, companies with above median ESG scores which improve their scores seem to be in particular demand from ESG investors, while companies with below ESG scores which improve their ESG scores are not systematically related to this demand. The demand seems to be related to the stock's expected return, confirming the implications of Fama and French (2007). However, the U-shape pattern of ESG and a stock's return (Kim and Statman 2012) is not observed. In sum, the sensitivity analysis in this subsection provides some additional support for the demand channel.

Alternative cutoffs

ESG scores are typically not continuously distributed, and the long/short portfolio in our base case consists of stocks with a long position in companies which have an ESG score equal to or above the median and a short position in stocks with an ESG score below the median. Thus, we allocate stocks with a median ESG stock into the long portfolio. This somewhat arbitrary allocation of median ESG stocks can be criticized (some difficulties are discussed, for example, in Gregory and Whittaker 2013). We therefore analyze how alternative cutoffs impact the results. We have changed the composition of the long and the short leg of the HL portfolio in four ways. First, we allocate companies with median ESG scores to the short leg. Second, we drop median ESG stocks from the allocation process. These two alternatives only change the results marginally—if at all—and are therefore not reported. Third, instead of buying (selling) stocks above

Table 4 The ESG return puzzle, cash-flow betas, and discount rate betas of the portfolio HL using alternative cutoffs

	50% (base case) (1)	40% (2)	30% (3)	20% (4)	10% (5)
Panel A: ESG return puzzle: cash-flow channel and discount rate channel					
R^{HL}	2.26%	3.72%	4.27%	4.91%	7.47%
ER^{HL}	−0.42%	−0.45%	−0.47%	−0.57%	−0.59%
UR^{HL}	2.68%	4.17%	4.74%	5.48%	8.06%
NCF^{HL}	0.25%	−0.95%	−1.59%	−3.16%	−3.38%
$-\text{NDR}^{\text{HL}}$	2.43%	5.12%	6.33%	8.63%	11.43%
Panel B: Risk mechanism					
β^{CF}	0.000	0.009	0.006	0.000	0.037
	(0.001)	(0.261)	(0.146)	(0.002)	(0.492)
$\beta^{\text{CF+}}$	−0.106	−0.153	−0.115	−0.142	−0.195
	(−0.001)	(−0.148)	(−0.077)	(−0.001)	(−0.258)
β^{DR}	−0.064	−0.072	−0.086	−0.106	−0.183
	(−4.441)	(−4.550)	(−4.873)	(−4.630)	(−5.341)
$\beta^{\text{DR+}}$	−0.106	−0.153	−0.115	−0.142	−0.195
	(−1.920)	(−2.079)	(−2.059)	(−2.389)	(−2.217)
Panel C: Demand mechanism					
β^{demand}	26.744	24.305	36.645	45.624	105.053
	(2.769)	(1.753)	(2.074)	(2.329)	(2.588)

This table reports average returns of the portfolio HL, which is long in H stocks (stocks with above or equal to median ESG scores) and short in L stocks with below median ESG scores). Returns of each in the portfolios H and L are equally weighted, the portfolio HL is long in H and short in L. Expected returns (ER) are derived from the four different models developed by Gebhardt et al. (2001), Claus and Thomas (2001), Ohlson and Juettner-Nauroth (2005), and Easton (2004), respectively. "Average" refers to the mean expected return across the four models. UR refers to the unexpected return which is decomposed into cash-flow news (NCF) and discount rate news (−NDR) using the method of Campbell and Shiller (1988). All returns are annualized using monthly returns in the calculations. Cash-flow and discount rate betas are computed following Campbell and Vuolteenaho (2004) as $\beta_t^{\text{CF}} \equiv \text{cov}_t\left(\text{UR}_{t+1}^{\text{HL}}, \text{NCF}_{t+1}^{M}\right)/\text{var}_t\left(\text{UR}_{t+1}^{M}\right)$ and $\beta_t^{\text{DR}} \equiv \text{cov}_t\left(\text{UR}_{t+1}^{\text{HL}}, -\text{NDR}_{t+1}^{M}\right)/\text{var}_t\left(\text{UR}_{t+1}^{M}\right)$, while the computation of the additional up betas follows Botshekan et al. (2012) as $\beta^{\text{CF+}} = \text{cov}\left(\text{UR}_{t+1}^{\text{HL}}, \text{NCF}_{t+1}^{M} | \text{UR}_{t+1}^{M} > 0\right)/\text{var}\left(\text{UR}_{t+1}^{M}\right)$ and $\beta^{\text{DR+}} = \text{cov}\left(\text{UR}_{t+1}^{\text{HL}}, -\text{NDR}_{t+1}^{M} | \text{UR}_{t+1}^{M} > 0\right)/\text{var}\left(\text{UR}_{t+1}^{M}\right)$
The demand beta β^{demand} is obtained from the regression $-\text{NDR}_t^{\text{HL}} = \alpha + \beta^{\text{demand}} \cdot \Delta\text{demand}_t + \varepsilon_t$ using quarterly returns over the sample period 2008 to 2018. t-values in parentheses

Table 5 The ESG return puzzle, cash-flow betas, and discount rate betas of the portfolio HL using alternative cutoffs

	Base case (cutoff = median) (1)	L = [0–4] H = [6–10] (2)	L = [0–3] H = [7–10] (3)
Panel A: ESG return puzzle: cash-flow channel and discount rate channel			
R^{HL}	2.26%	3.67%	5.92%
$\mathrm{ER}^{\mathrm{HL}}$	−0.42%	−0.58%	−0.51%
$\mathrm{UR}^{\mathrm{HL}}$	2.68%	4.24%	6.43%
$\mathrm{NCF}^{\mathrm{HL}}$	0.25%	−2.17%	−2.89%
$-\mathrm{NDR}^{\mathrm{HL}}$	2.43%	6.41%	9.32%
Panel B: Risk mechanism			
β^{CF}	0.000	0.002	0.029
	(0.001)	(0.038)	(0.451)
$\beta^{\mathrm{CF+}}$	−0.106	−0.150	−0.146
	(−0.001)	(−0.021)	(−0.232)
β^{DR}	−0.064	−0.117	−0.105
	(−4.441)	(−6.731)	(−3.370)
$\beta^{\mathrm{DR+}}$	−0.106	−0.150	−0.176
	(−1.920)	(−2.870)	(−1.424)
Panel C: Demand mechanism			
β^{demand}	26.744	38.491	95.274
	(2.769)	(2.334)	(2.857)

This table reports average returns of the portfolio HL which is long in H stocks (stocks with above or equal to median ESG scores) and short in L stocks with below median ESG scores). Returns of each in the portfolios H and L are equally weighted; the portfolio HL is long in H and short in L. Expected returns (ER) are derived from the four different models developed by Gebhardt et al. (2001), Claus and Thomas (2001), Ohlson and Juettner-Nauroth (2005), and Easton (2004), respectively "Average" refers to the mean expected return across the four models. UR refers to the unexpected return, which is decomposed into cash-flow news (NCF) and discount rate news (−NDR) using the method of Campbell and Shiller (1988). All returns are annualized using monthly returns in the calculations. Cash-flow and discount rate betas are computed following Campbell and Vuolteenaho (2004) as $\beta_t^{\mathrm{CF}} \equiv \mathrm{cov}_t\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, \mathrm{NCF}_{t+1}^{M}\right)/\mathrm{var}_t\left(\mathrm{UR}_{t+1}^{M}\right)$ and $\beta_t^{\mathrm{DR}} \equiv \mathrm{cov}_t\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, -\mathrm{NDR}_{t+1}^{M}\right)/\mathrm{var}_t\left(\mathrm{UR}_{t+1}^{M}\right)$, while the computation of the additional up betas follows Botshekan et al. (2012) as $\beta^{\mathrm{CF+}} = \mathrm{cov}\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, \mathrm{NCF}_{t+1}^{M} | \mathrm{UR}_{t+1}^{M} > 0\right)/\mathrm{var}\left(\mathrm{UR}_{t+1}^{M}\right)$ and $\beta^{\mathrm{DR+}} = \mathrm{cov}\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, -\mathrm{NDR}_{t+1}^{M} | \mathrm{UR}_{t+1}^{M} > 0\right)/\mathrm{var}\left(\mathrm{UR}_{t+1}^{M}\right)$
The demand beta β^{demand} is obtained from the regression $-\mathrm{NDR}_t^{\mathrm{HL}} = \alpha + \beta^{\mathrm{demand}} \cdot \Delta\mathrm{demand}_t + \varepsilon_t$ using quarterly returns over the sample period 2008 to 2018. t-values in parentheses

(below) the median ESG score, we use alternative cutoffs. In particular, we sell those stocks with an ESG score below the p-quantile and we buy stocks with an ESG score above the $(1-p)$-quantile. We set p to 10%, 20%, 30%, and 40% (see Table 4). Finally, we use the level of the aggregate ESG score as a cutoff criterion. We then form the HL portfolio using the following cutoffs: short leg: ESG score below or equal to 3 (4); long leg: ESG score above or equal to 7 (6) (see Table 5).

In general, the results remain largely consistent with those presented earlier. They support the existence of the ESG return puzzle and its explanation by the demand channel. Furthermore, they are consistent with a further economic implication of the Fama and French (2007) demand model. That is, a tighter cutoff leads to a larger unexpected return, since stocks with a higher ESG score should experience a higher demand from investors with ESG preferences. For example, if the long (short) portfolio contains just 10% stocks with the highest (lowest) ESG score, the unexpected return increases from 2.68% (median cutoff) to 8.06% (Panel A). Thereby, the unexpected return can be primarily attributed to discount rate news (increasing from 2.43 to 11.43%). However, the risk channel does not seem to be supported by the data, since the decomposed discount rate betas are negative for all cutoffs (Panel B) and in most cases even significantly smaller than zero. Thus, there is no support for the hypothesis that H stocks are riskier than L stocks are. Rather, we observe that discount rate betas tend to become more negative for a tighter cutoff. For example, the discount rate beta falls from −0.064 (column (1)) to −0.183 (column (5)), indicating that H stocks become even less risky than L stocks. Further, the demand beta increases with a tighter cutoff (see Panel C in Table 4), which is an implication of the demand model of Fama and French (2007). However, we also observe a tendency that demand betas are estimated with lower precision when applying a tighter cutoff. This observation can be attributed to the fact that a tighter cutoff reduces the number of stocks in the HL portfolio, which makes the point estimate of the demand beta less precise.

Similar results are observed when we use absolute values of the ESG score as cutoffs: a tighter cutoff leads to a larger unexpected return (see Table 5). For example, if the long (short) portfolio contains those stocks with an ESG score above or equal to 7 (below or equal to 3), see last column, the unexpected return equals 6.43% per annum. The decomposition of this unexpected return reveals that cash-flow news contributes negatively (−2.89%) while discount rate news contributes positively (9.32%). These observations confirm our previous conclusions that discount rates of high-ESG companies have fallen to a larger extent than those companies with low ESG ratings. Thereby, the risk channel is unable to explain the changes in the discount rate, as we observe a negative discount rate beta of the HL portfolio. In contrast, the demand channel receives additional support. That is, the share of ESG investors seems to be significantly related to the unexpected changes in discount rates. The estimated slope coefficient is about 95. In economic terms this coefficient implies that an increase in the ratio of ESG investors to all investors by 10%-points increases prices of ESG companies with a high ESG rating by 9.5% relative to very low-ESG companies. The results are consistent with the view that prices of H stocks rise faster (relative to L stocks) when demand from ESG investors increases and the rising prices can be explained by lower discount rates, which, however, are not driven by risk characteristics. In sum, alternative cutoffs are consistent with the demand channel.

Theme-specific ratings

The previous sections have analyzed the aggregate ESG rating of MSCI, which is an industry-adjusted rating. Galema et al. (2008) argue that an aggregation over different ESG dimensions may have confounding effects and potentially introduce errors into the analyses. In this subsection, we therefore focus on the different dimensions of ESG individually, namely the ecological (E), social (S), and governance (G) dimensions. Although the MSCI rating methodology would allow us to use even more detailed dimensions, we abstain from doing so for several reason. First, it is likely that a more detailed level of the various dimensions of ESG is associated with a larger measurement error. Second, if ESG ratings of companies are made available to the public, it is mainly the top level rating. Thus, a more detailed level of ESG ratings is less likely to be recognized by investors.

The return decomposition is different for E, S, and G (see Table 6). Although for each score the unexpected return is positive (confirming the ESG return puzzle), its decomposition delivers alternative explanations. Cash-flow news is somewhat positive for S and G, while discount rate news of the HL portfolio using only S scores is negative. Also, the unexpected return is the lowest for the HL portfolio using solely the S score.[7] Adding the absolute scores of E, S, and G (denoted by E+S+G) and forming a HL portfolio results in an unexpected return of 3.56% per annum, outperforming the HL portfolio using the best-in-class ESG score (base case, see last column). The higher unexpected return of about 1% per annum, however, can be attributed to better cash-flow news. Thus, an ESG score using an absolute approach seems to select stocks that improve their fundamentals more effectively than using the best-in-class ESG scoring does.

The risk channel is not able to explain the decomposition results. Looking at the cash-flow and discount rate betas in Panel B, none of them is significantly larger than zero. The demand beta, displayed in Panel C, however, is larger than zero although the S dimension does not produce a significant coefficient (see column (2)). In sum, the analysis of discount rate news in relation to single E, S, and G measures suggests that the variation in expected returns is primarily driven by changes in investor demand for such characteristics.

[7] Brammer et al. (2006) even find a negative relation between an S score and the stock return in the UK.

Table 6 The ESG return puzzle, cash-flow betas, and discount rate betas of the portfolio HL using individual E, S, and G scores

	E (1)	S (2)	G (3)	E+S+G (4)	ESG (5)
Panel A: ESG return puzzle: cash-flow channel and discount rate channel					
R^{HL}	3.23%	0.82%	3.14%	3.31%	2.26%
ER^{HL}	−0.23%	−0.29%	−0.15%	−0.25%	−0.42%
UR^{HL}	3.45%	1.12%	3.28%	3.56%	2.68%
NCF^{HL}	−0.95%	1.69%	0.84%	1.51%	0.25%
$-NDR^{HL}$	4.40%	-0.58%	2.44%	2.05%	2.43%
Panel B: Risk mechanism					
β^{CF}	−0.028	−0.001	0.000	−0.019	0.000
	(−0.778)	(−0.026)	(0.010)	(−0.583)	(0.001)
β^{CF+}	0.031	−0.014	0.118	−0.046	−0.106
	(0.463)	(−0.015)	(0.006)	(−0.356)	(−0.001)
β^{DR}	−0.024	−0.027	−0.012	−0.026	−0.064
	(−1.321)	(−2.015)	(−0.679)	(−1.627)	(−4.441)
β^{DR+}	−0.031	−0.014	0.118	−0.046	−0.106
	(−0.663)	(−0.978)	(0.335)	(−0.836)	(−1.920)
Panel C: Demand mechanism					
β^{demand}	41.396	3.588	54.548	27.531	26.744
	(2.852)	(0.288)	(4.374)	(1.937)	(2.769)

This table reports average returns of the portfolio HL, which is long in H stocks (stocks with above or equal to median ESG scores) and short in L stocks with below median ESG scores). Returns of each in the portfolios H and L are equally weighted; the portfolio HL is long in H and short in L. Expected returns (ER) are derived from the four different models developed by Gebhardt et al. (2001), Claus and Thomas (2001), Ohlson and Juettner-Nauroth (2005), and Easton (2004), respectively. "Average" refers to the mean expected return across the four models. UR refers to the unexpected return, which is decomposed into cash-flow news (NCF) and discount rate news (−NDR) using the method of Campbell and Shiller (1988). All returns are annualized using monthly returns in the calculations. Cash-flow and discount rate betas are computed following Campbell and Vuolteenaho (2004) as $\beta_t^{CF} \equiv \mathrm{cov}_t\left(UR_{t+1}^{HL}, NCF_{t+1}^{M}\right)/\mathrm{var}_t\left(UR_{t+1}^{M}\right)$ and $\beta_t^{DR} \equiv \mathrm{cov}_t\left(UR_{t+1}^{HL}, -NDR_{t+1}^{M}\right)/\mathrm{var}_t\left(UR_{t+1}^{M}\right)$, while the computation of the additional up betas follows Botshekan et al. (2012) as $\beta^{CF+} = \mathrm{cov}\left(UR_{t+1}^{HL}, NCF_{t+1}^{M} | UR_{t+1}^{M} > 0\right)/\mathrm{var}\left(UR_{t+1}^{M}\right)$ and $\beta^{DR+} = \mathrm{cov}\left(UR_{t+1}^{HL}, -NDR_{t+1}^{M} | UR_{t+1}^{M} > 0\right)/\mathrm{var}\left(UR_{t+1}^{M}\right)$
The demand beta β^{demand} is obtained from the regression $-NDR_t^{HL} = \alpha + \beta^{demand} \cdot \Delta demand_t + \varepsilon_t$ using quarterly returns over the sample period 2008 to 2018. *t*-values in parentheses

Alternative weighting approaches and time periods

The base case approach uses an equal-weighted H and L portfolio to obtain the HL portfolio. In this subsection, we use a value-weighted HL portfolio. Additionally, we limit the sample to the most recent period 2013 to 2018 (instead of the full sample period from 2008 to 2018). Table 7 summarizes the results for these alternatives. Column (1) displays the base case approach (equal-weighted HL portfolio), column (2) shows the results for the value-weighted HL

Table 7 The ESG return puzzle, cash-flow betas, and discount rate betas of the portfolio HL using value—weights and the recent sample period

	Equal-weighted portfolios (2008–2018)—(base case) (1)	Value-weighted portfolios (2008–2018) (2)	Value-weighted portfolios (2013–2018) (3)
Panel A: ESG return puzzle: cash-flow channel and discount rate channel			
R^{HL}	2.26%	3.31%	4.32%
$\mathrm{ER}^{\mathrm{HL}}$	−0.42%	−0.25%	−0.19%
$\mathrm{UR}^{\mathrm{HL}}$	2.68%	3.56%	4.51%
$\mathrm{NCF}^{\mathrm{HL}}$	0.25%	1.51%	−0.82%
$-\mathrm{NDR}^{\mathrm{HL}}$	2.43%	2.05%	5.32%
Panel B: Risk mechanism			
β^{CF}	0.000	−0.032	−0.001
	(0.001)	(−0.711)	(−0.041)
$\beta^{\mathrm{CF+}}$	−0.106	0.034	−0.052
	(−0.001)	(0.412)	(−0.027)
β^{DR}	−0.064	−0.131	−0.004
	(−4.441)	(−5.850)	(−0.349)
$\beta^{\mathrm{DR+}}$	−0.106	−0.034	−0.052
	(−1.920)	(−2.860)	(−0.195)
Panel C: Demand mechanism			
β^{demand}	26.744	27.531	126.105
	(2.769)	(1.937)	(2.581)

This table reports average returns of the portfolio HL, which is long in H stocks (stocks with above or equal to median ESG scores) and short in L stocks with below median ESG scores). Returns of each in the portfolios H and L are equally weighted; the portfolio HL is long in H and short in L. Expected returns (ER) are derived from the four different models developed by Gebhardt et al. (2001), Claus and Thomas (2001), Ohlson and Juettner-Nauroth (2005), and Easton (2004), respectively. "Average" refers to the mean expected return across the four models. UR refers to the unexpected return which is decomposed into cash-flow news (NCF) and discount rate news (−NDR) using the method of Campbell and Shiller (1988). All returns are annualized using monthly returns in the calculations. Cash-flow and discount rate betas are computed following Campbell and Vuolteenaho (2004) as $\beta_t^{\mathrm{CF}} \equiv \mathrm{cov}_t\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, \mathrm{NCF}_{t+1}^{M}\right)/\mathrm{var}_t\left(\mathrm{UR}_{t+1}^{M}\right)$ and $\beta_t^{\mathrm{DR}} \equiv \mathrm{cov}_t\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, -\mathrm{NDR}_{t+1}^{M}\right)/\mathrm{var}_t\left(\mathrm{UR}_{t+1}^{M}\right)$, while the computation of the additional up betas follows Botshekan et al. (2012) as $\beta^{\mathrm{CF+}} = \mathrm{cov}\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, \mathrm{NCF}_{t+1}^{M} | \mathrm{UR}_{t+1}^{M} > 0\right)/\mathrm{var}\left(\mathrm{UR}_{t+1}^{M}\right)$ and $\beta^{\mathrm{DR+}} = \mathrm{cov}\left(\mathrm{UR}_{t+1}^{\mathrm{HL}}, -\mathrm{NDR}_{t+1}^{M} | \mathrm{UR}_{t+1}^{M} > 0\right)/\mathrm{var}\left(\mathrm{UR}_{t+1}^{M}\right)$

The demand beta β^{demand} is obtained from the regression $-\mathrm{NDR}_t^{\mathrm{HL}} = \alpha + \beta^{\mathrm{demand}} \cdot \Delta\mathrm{demand}_t + \varepsilon_t$ using quarterly returns over the sample period 2008 to 2018. t-values in parentheses

portfolio, and column (3) reports return characteristics for the more recent period.

Looking at the ESG return puzzle displayed in Panel A, we find that the unexpected return of the value-weighted HL portfolio over the full sample (column (2)) and the more recent sample (column (3)) is even larger than in the base case (3.56% and 4.51% compared to 2.68%). Thus, the base case seems to be a conservative approach to estimating the size of the ESG return puzzle. The unexpected return of the HL portfolio is driven primarily by discount rate news. Looking at the risk mechanism in Panel B and the demand mechanism in Panel C, the value-weighted HL portfolio in the full sample and the more recent sample largely confirm our previous conclusions. We find no evidence that stocks with a good ESG rating are more risky than those with a bad ESG rating. If at all, betas are smaller for H companies than for L companies, indicating that good-ESG firms are less risky than bad ESG firms are. However, the demand mechanism (shown in Panel C) receives support from the value-weighted HL portfolio, both in the full sample and the more recent period. It is interesting to note that the demand beta in the recent period is substantially larger than the estimate we obtain in the full sample (126 versus 27). Therefore, the demand model of Fama and French (2007) seems to be a particular good explanation of stock returns when looking at the last years compared with the base case. In sum, alternative weighting approaches and different sample periods support the previous conclusion about the ESG return puzzle and its demand explanation.

Discussion and conclusion

Empirically, ESG stocks in the USA have performed better than non-ESG stocks, although their expected returns are implied to be lower by valuation models. We refer to the different signs in realized and expected returns as the "ESG return puzzle." A decomposition analysis of the unexpected return (i.e., the difference between the realized and the expected returns) reveals that the ESG return puzzle can be primarily explained by discount rate news. That is, good-ESG stocks have performed better than bad ESG stocks because investors have applied a relatively smaller discount rate to the former. Cash-flow news, however, is not systematically linked to the performance of ESG stocks relative to non-ESG stocks. This main result is robust to the identification of cash-flow news and discount rate news for which we use four different models.

We investigate two approaches which potentially explain the discount rate channel of the ESG return puzzle, i.e., the risk mechanism and the demand mechanism. We find that higher realized returns of ESG stocks (compared to non-ESG stocks) cannot be explained by higher discount rate betas. This implies that traditional risk measures provide no explanation for the ESG return puzzle. We also investigate the demand channel, which is based on the model more recently proposed by Fama and French (2007). In their model, the demand for assets with non-financial characteristics such as ESG are an important driver of expected returns. Approximating such an ESG demand by the ratio of ESG investors relative to all investors helps to explain discount rate news of good-ESG companies relative to bad ESG companies. If the demand of investors with ESG preferences increases, it drives prices of good-ESG companies upward and expected returns downward. We provide several sensitivity analyses which support the existence of the ESG return puzzle and its explanation through the demand mechanism. The observation that ESG stocks are primarily driven by investor demand (and not by improving cash-flow prospects or improving risk characteristics) has several implications for investors, policy makers, and companies alike.

Investors should be aware that if their demand drives stock prices up and discount rates (expected returns) down, past returns are a bad guide for future returns. In this case, higher realized returns for good-ESG companies relative to bad ESG companies can only be extrapolated into the future if additional demand from new investors with ESG preferences hits the market. However, at some point in time, the ratio of ESG investors cannot increase further (because 100% of investors have ESG preferences or because there is a stable equilibrium between ESG and non-ESG investors). Then, theory implies that realized returns should equal their expected returns over the long run. If the capital market reaches this point, ESG assets should deliver returns that are lower than in the past. This may disappoint some investors in the long-term if they are not willing to accept lower returns for holding ESG assets.

Second, from the perspective of the economy and policy makers, such a point may be desirable. Companies with good ESG characteristics can exploit a lower cost of capital relative to companies with bad ESG characteristics; thus, they have a competitive advantage and can finance their investments at lower costs. In the long term, the economy will improve their ESG characteristics. If this is the intention of policy makers, an additional regulatory framework, such as the European Union's Taxonomy (European Union 2019), which is currently under discussion and which will most likely be implemented in the near future, will support and increase the speed of the demand channel. Then, capital markets will become an effective tool for implementing ESG policies. What is important to note is that such a mechanism works independently of the common risk-return relations underlying most asset pricing models.

Third, companies should consider the key contribution of this paper: that a substantial part of the variation in expected returns (cost of capital) is explained by a demand from ESG investors which is not related to risk. The management of a company should be aware that there is such a preference function of investors which is partly unrelated to financial issues. This preference function seems to have changed toward ESG, and managers of firms that have recognized this seem to have profited from investor demand through a reduction in the company's cost of equity capital which has led to a higher share price. This relation highlights a strategic management issue: that of knowing the preferences of potential or actual shareholders. The question arises of who should earn the benefits from a reduced cost of capital and higher stock prices. The answer should concern shareholders when they set appropriate incentives for their management and appropriate rules for their remuneration. If these incentives are (partly) related to the stock price, a good incentive system should differentiate between an increase in the share price stemming from good cash flows, from risk-reducing strategies, and from higher demand from investors (e.g., investors with non-financial preferences). Therefore, the results of this study may provide managers and shareholders alike with a roadmap on how non-financial characteristics, such as ESG, relate to financial returns in the long-term perspective.

While our analysis is restricted to US stocks, we expect that results in other stock markets should be similar to those reported in this paper. The main justification for this conjecture is the observation that the trend toward ESG investing is a global trend not only limited to the USA. The results of our study are derived from using the MSCI ESG rating methodology. While MSCI ratings are one of the important sustainability ratings, it remains an open issue left for future research whether the same conclusions can be drawn from alternative ratings. However, the analysis of the drivers of returns (i.e., cash-flow news and discount rate news) and their potential mechanism (risk and demand) adds value to the understanding of the return differences between high- and low-ESG stocks. A focus on only returns may not be able to distinguish between the two explanations. In any case, the results provide investors and corporate managers with more complete information about how ESG relates to returns (expected and realized).

Acknowledgements The author acknowledges the access to ESG data provided by Quoniam Asset Management.

References

Albuquerque, R., Y. Koskinen, and C. Zhang. 2019. Corporate social responsibility and firm risk: Theory and empirical evidence. *Management Science* 65:4451–4469.

Armstrong, J.S., and K.C. Green. 2013. Effects of corporate social responsibility and irresponsibility policies. *Journal of Business Research* 66:1922–1927.

Barnett, M.L. 2007. Stakeholder influence capacity and the variability of financial returns to corporate social responsibility. *Academy of Management Review* 32:794–816.

Barney, J. 1991. Firm resources and sustained competitive advantage. *Journal of Management* 33:1304–1320.

Bauer, R., N. Gunster, and R. Otten. 2004. Empirical evidence on corporate governance in Europe. *Journal of Asset Management* 5:91–104.

Berg, F.; J. Koelbel, and R. Rigobon. 2019. Aggregate confusion: The Divergence of ESG ratings. MIT Sloan Working Paper 5822-19. MIT Sloan School of Management (August).

Bolton, P., and M. Kacperczyk. 2020. Carbon premium around the world. Working Paper Columbia University.

Botshekan, M., R. Kraeussl, and A. Lucas. 2012. Cash-flow and discount rate risk in up and down markets: What is actually priced?. *Journal of Financial and Quantitative Analysis* 47:1279–1301.

Brzeszczynski, J., and G. McIntosh. 2014. Performance of portfolios composed of British SRI stocks. *Journal of Business Ethics* 120:335–362.

Campbell, J.Y., and R.J. Shiller. 1988. Stock prices, earnings, and expected dividends. *Journal of Finance* 43:661–676.

Campbell, J.Y., C. Polk, and T. Vuolteenaho. 2010. Growth or glamour? Fundamentals and systematic risk in stock returns. *Review of Financial Studies* 23:305–344.

Campbell, J.Y., and T. Vuolteenaho. 2004. Bad beta, good beta. *American Economic Review* 94:1249–1275.

Carhart, M.M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52:57–82.

Chava, S. 2014. Environmental externalities and cost of capital. *Management Science* 60 (9):2223–2247.

Chen, L., Z. Da, and X. Zhao. 2013. What drives stock price movements?. *Review of Financial Studies* 26:841–876.

Chen, L., and X. Zhao. 2009. Return decomposition. *Review of Financial Studies* 22:5213–5249.

Choi, D., Z. Gao, and W. Jiang. 2020. Attention to global warming. *Review of Financial Studies* 33:1112–1145.

Choi, J., and H. Wang. 2009. Stakeholder relations and the persistence of corporate social performance. *Strategic Management Journal* 30:895–907.

Claus, J., and J. Thomas. 2001. Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets. *Journal of Finance* 56:1629–1666.

Colbert, B.A. 2004. The complex resource-based view: Implications for theory and practice in strategic human resource management. *Academy of Management Review* 29:341–358.

Dhaliwal, D.S., O.Z. Li, A. Tsang, and Y.G. Yang. 2020. Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *Accounting Review* 86:59–100.

Easton, P.D. 2004. PE ratios, PEG ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review* 79:73–95.

Eccles, R.G., I. Ioannou, and G. Serafeim. 2014. The impact of corporate sustainability on organizational processes and performance. *Management Science* 60:2835–2857.

El Ghoul, S., O. Guedhami, H. Kim, and K. Park. 2018. Corporate environmental responsibility and the cost of capital: International evidence. *Journal of Business Ethics* 149:335–361.

El Ghoul, S., O. Guedhami, C.C.Y. Kwok, and D.R. Mishra. 2011. Does corporate responsibility affect the cost of capital?. *Journal of Banking and Finance* 35:2388–2406.

EU. 2019. EU technical expert group on sustainable finance: Financing a sustainable European economy. Taxonomy technical report.

Extel Survey. 2019. Independent Research in Responsible Investment Survey 2019.

Fama, E.F. 1991. Efficient capital markets: II. *Journal of Finance* 46:1575–1617.

Fama, E.F., and K.R. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33:3–56.

Fama, E.F., and K.R. French. 2007. Disagreement, tastes, and asset prices. *Journal of Financial Economics* 83:667–689.

Fama, E.F., and K.R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics.* 116:1–22.

Fama, E.F., and K.R. French. 2018. Choosing factors. *Journal of Financial Economics* 128:234–252.

Financial Times. 2018. BlackRock stakes claim on 'sustainable investing' revolution. *FTfm, 19 Nov 2018, page 6.*

Friedman, M. 1970. The social responsibility of business is to increase profits. *Times Magazine*, New York, September 13, 32–33.

Galema, R., A. Plantiga, and B. Scholtens. 2008. The stocks at stake: Return and risk in socially responsible investment. *Journal of Banking and Finance* 32:2646–2654.

Gebhardt, W., C. Lee, and B. Swaminathan. 2001. Towards an implied cost of capital. *Journal of Accounting Research* 39:135–176.

Godfrey, P., C. Merrill, and J. Hansen. 2009. The relationship between CSR and shareholder value: An empirical test of the risk management hypothesis. *Strategic Management Journal* 30:425–455.

Gregory, A., R. Tharyan, and J. Whittaker. 2014. Corporate social responsibility and firm value: Disaggregating the effects on cash-flow, risk and growth. *Journal of Business Ethics* 124:633–657.

Gregory, A., and J. Whittaker. 2013. Exploring the valuation of corporate social responsibility—A comparison of research methods. *Journal of Business Ethics* 116:1–20.

Gregory, R.P., J.G. Stead, and E. Stead. 2020. The global pricing of environmental, social, and governance (ESG) criteria. *Journal of Sustainable Finance & Investment* 26:1–20.

Hong, B., Z. Li, and D. Minor. 2016. Corporate governance and executive compensation for corporate social responsibility. *Journal of Business Ethics* 136:199–213.

Hong, H., and K. Kacperczyk. 2009. The price of sin: The effects of social norms on markets. *Journal of Financial Economics* 93:15–36.
Ilhan, E., Z. Sautner, and G. Vilkov. 2020. Carbon tail risk. *Review of Financial Studies*, forthcoming.
Jensen, M. 2002. Value maximization, stakeholder theory, and the corporate objective function. *Business Ethics Quarterly* 12:235–256.
Kempf, A., and P. Osthoff. 2007. The effect of socially responsible investing on portfolio performance. *European Financial Management* 13:908–922.
Kim, Y., and M. Statman. 2012. Do corporations invest enough in environmental responsibility?. *Journal of Business Ethics* 105:115–129.
Li, F., and A. Polychronopoulos. 2020. What a difference an ESG ratings provider makes! Research Affiliates, January 2020.
Li, Z.F., D.B. Minor, J. Wang, and C. Yu. 2019. A learning curve of the market: Chasing alpha of socially responsible firms. *Journal of Economic Dynamics and Control* 109:1–17.
Liang, H., and L. Renneboog. 2020. Corporate social responsibility and sustainable finance: A review of the literature. ECGI Working Paper Series in Finance 701/2020.
Lintner, J. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47:13–37.
Luo, H.A., and R.J. Balvers. 2017. Social screens and systematic investor boycott risk. *Journal of Financial and Quantitative Analysis* 52:365–399.
McWilliam, A., and D.S. Siegel. 2011. Creating and capturing value: Strategic corporate social responsibility, resource-based theory, and sustainable competitive advantage. *Journal of Management* 37:1480–1495.
Merton, R.C. 1973. An intertemporal capital asset pricing model. *Econometrica* 41:867–887.
Oehmke, M., and M. Opp. 2020. A theory of socially responsible investment. Working Paper.
Ohlson, J., and B. Juettner-Nauroth. 2005. Expected EPS and EPS growth as determinants of value. *Review of Accounting Studies* 10:349–365.
Pastor, L., R.F. Stambaugh, and L.A. Taylor. 2020. Sustainable investing in equilibrium. *Journal of Financial Economic*, forthcoming.
Riedl, A., and P. Smeets. 2017. Why do investors hold socially responsible mutual funds. *Journal of Finance* 72:2505–2550.
Revelli, C., and J.-L. Viviani. 2015. Financial performance of socially responsible investing (SRI): What have we learned? A meta-analysis. *Business Ethics: A European Review* 24:158–185.
Robinson, M., A. Kleffner, and S. Bertels. 2011. Signaling sustainability leadership: Empirical evidence of the value of DJSI membership. *Journal of Business Ethics* 101:493–505.
Sen, S., and C.B. Bhattachary. 2001. Does doing good always lead to doing better? Consumer reactions to corporate social responsibility. *Journal of Marketing Research* 38:225–243.
Sharpe, W.F. 1964. Capital asset prices: A theory of market equilibrium under conditions of Risk. *Journal of Finance* 19:425–442.
SustainAbility (2020). Rate the raters 2020: Investor survey and interview results. March 2020.
Tang, Y., J.G. Wu, and L. Zhang. 2014. Do anomalies exist ex ante?. *Review of Finance* 18:843–875.
Van Duuren, A., A. Plantinga, and B. Scholtens. 2016. ESG integration and the investment management process: Fundamental investing reinvented. *Journal of Business Ethics* 138:525–533.
Xiao, Y., R. Faff, P. Gharghori, and D. Lee. 2013. An empirical study of the world price of sustainability. *Journal of Business Ethics* 114:297–310.
Zerbib, O.D. 2019. The effect of pro-environmental preferences on bond prices: Evidence from green bonds. *Journal of Banking and Finance* 98:39–60.
Zerbib, O.D. 2020. A sustainable asset pricing model (S-CAPM): Evidence from green investing and sin stock exclusion. Working Paper, Tilburg University.

Dr. Olaf Stotz is Full Professor of Asset Management and Pension Economics at Frankfurt School of Finance & Management. He received his PhD degree and Habilitation degree at the University of Aachen. His research interests include asset management, behavioral and empirical finance, and pension economics and retirement decisions. He has extensive academic and professional experience in those fields, working at leading Universities and investment companies. He serves on various expert panels and supervisory boards.