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

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

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

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

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

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

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^{HL}_{FF3}	3.05	3.11	2.61	2.28	2.76
R^{HL}_{C4}	3.06	3.11	2.61	2.29	2.76
R^{HL}_{FF5}	2.64	2.67	2.22	1.89	2.34
R^{HL}_{FF6}	2.63	2.66	2.21	1.89	2.32
Panel C: Decomposed average returns of portfolio HL					
ER^{HL}	−0.42	−0.14	−0.35	−0.60	−0.36
UR^{HL}	2.68	2.45	2.21	2.23	2.34
NCF^{HL}	0.25	−1.14	−1.25	−0.09	0.23
$-\mathrm{NDR}^{HL}$	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 $\overline{R^{HL}}$. 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, $\overline{R^{H}} > \overline{R^{L}}$. 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, $\overline{R^{HL}}$, 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 ($\overline{R^{HL}} > 0$).

[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 $E(R^{\mathrm{HL}})$, 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 ($E(R^{\mathrm{HL}}) > 0$), 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 $E(R^{\mathrm{HL}})$ 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 $\beta_t^{CF} \equiv cov_t\left(UR_{t+1}^{HL}, NCF_{t+1}^{M}\right)/var_t\left(UR_{t+1}^{M}\right)$ and $\beta_t^{DR} \equiv cov_t\left(UR_{t+1}^{HL}, -NDR_{t+1}^{M}\right)/var_t\left(UR_{t+1}^{M}\right)$, while the computation of the additional up betas follows Botshekan et al. (2012) as $\beta^{CF+} = cov\left(UR_{t+1}^{HL}, NCF_{t+1}^{M} | UR_{t+1}^{M} > 0\right)/var\left(UR_{t+1}^{M}\right)$ and $\beta^{DR+} = cov\left(UR_{t+1}^{HL}, -NDR_{t+1}^{M} | UR_{t+1}^{M} > 0\right)/var\left(UR_{t+1}^{M}\right)$

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%
ER^{HL}	−0.42%	−0.58%	−0.51%
UR^{HL}	2.68%	4.24%	6.43%
NCF^{HL}	0.25%	−2.17%	−2.89%
$-\mathrm{NDR}^{HL}$	2.43%	6.41%	9.32%
Panel B: Risk mechanism			
β^{CF}	0.000	0.002	0.029
	(0.001)	(0.038)	(0.451)
β^{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)
β^{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 $\beta_t^{CF} \equiv \mathrm{cov}_t\left(\mathrm{UR}_{t+1}^{HL}, \mathrm{NCF}_{t+1}^{M}\right) / \mathrm{var}_t\left(\mathrm{UR}_{t+1}^{M}\right)$ and $\beta_t^{DR} \equiv \mathrm{cov}_t\left(\mathrm{UR}_{t+1}^{HL}, -\mathrm{NDR}_{t+1}^{M}\right) / \mathrm{var}_t\left(\mathrm{UR}_{t+1}^{M}\right)$, while the computation of the additional up betas follows Botshekan et al. (2012) as $\beta^{CF+} = \mathrm{cov}\left(\mathrm{UR}_{t+1}^{HL}, \mathrm{NCF}_{t+1}^{M} | \mathrm{UR}_{t+1}^{M} > 0\right) / \mathrm{var}\left(\mathrm{UR}_{t+1}^{M}\right)$ and $\beta^{DR+} = \mathrm{cov}\left(\mathrm{UR}_{t+1}^{HL}, -\mathrm{NDR}_{t+1}^{M} | \mathrm{UR}_{t+1}^{M} > 0\right) / \mathrm{var}\left(\mathrm{UR}_{t+1}^{M}\right)$ The demand beta β^{demand} is obtained from the regression $-\mathrm{NDR}_t^{HL} = \alpha + \beta^{demand} \cdot \Delta\mathrm{demand}_t + \varepsilon_t$ 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.

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 $-NDR_t^{HL} = \alpha + \beta^{demand} \cdot \Delta\mathrm{demand}_t + \varepsilon_t$ 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

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

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.