Patience pays off – corporate social responsibility and long-term stock returns

Gregor Dorfleitner [a], Sebastian Utz [b] and Maximilian Wimmer [a,c]

[a]Department of Finance, University of Regensburg, Regensburg, Germany; [b]School of Finance, Universität St.Gallen (HSG), St. Gallen, Switzerland; [c]Finance Area, University of Mannheim, Mannheim, Germany

ABSTRACT

This paper presents new evidence on the implications of corporate social responsibility (CSR) on stock returns. By implementing a long-term focus as well as using subdivided measures for CSR, we cater to the intangible nature and the heterogeneity of CSR activities. We use a novel classification of these activities into nine *areas*, each belonging to one of the standard environment, social, and governance (ESG) dimensions. Using cross-sectional return regressions and buy-and-hold abnormal returns, we find that firms with strong CSR significantly outperform firms with weak CSR in the mid and long run in certain areas. Firm returns increase up to 3.8% with respect to a one-standard-deviation increase of the CSR rating. In a two-stage least squares (2SLS) approach we verify that the main economic channel for the appreciation of strong CSR stocks is unexpected additional cash flows. The results are relevant for assessing the efficiency of CSR, and have broader implications for asset managers who can expect abnormal returns by investing in firms that exhibit a high CSR in the respective scores and holding the stocks for a longer period.

KEYWORDS

Asset pricing; corporate social responsibility; market efficiency; buy-and-Hold abnormal return; earnings surprises

JEL

G14; G30; G11; Q56

1. Introduction

Market efficiency in pricing corporate social responsibility (CSR) investments is subject of debate. A well-established stream of empirical work considers the implications of CSR on short-run stock returns (Gompers, Ishill, and Metrick 2003; Bauer, Guenster, and Otten 2004; Derwall et al. 2005; Statman and Glushkov 2009; Humphrey, Lee, and Shen 2012). However, CSR investments often yield intangible assets whose value is rather opaque to the markets until they start to generate tangible outcomes several years later (Edmans 2011). Therefore, this paper provides insights into the *long-term* effects of CSR activities on stock returns. With respect to these CSR activities, we use a novel classification of nine different fields of CSR activities to capture the entire CSR of firms. Our central strategy is to analyze the time structure of abnormal returns of a cross-section of US and Canadian stocks dependent on CSR, which we measure with ESG scores – an acronym for environment, social, and (corporate) governance scores. We find positive mid- and long-term effects in the environment and social dimensions of up to a 3.8%

CONTACT Gregor Dorfleitner gregor.dorfleitner@ur.de Department of Finance, University of Regensburg, 93040 Regensburg, Germany

abnormal return with respect to a one-standard-deviation change of the corresponding ESG score. We identify *emission and resource reduction*, *workforce*, and *society* as exactly the areas in which CSR investments are profitable. The abnormal returns are robust to controlling for common risk factors, accounting data, as well as year and industry-fixed effects. Finally, we verify that the economic channel for the appreciation of strong CSR stocks is mainly unexpected additional cash flows for the respective stocks during our sample period of 2002–2014.

Since there are relatively few natural experiments that result in significant shocks in CSR, most studies rely on exogenous information proxies for the level of CSR of single firms (Cheng, Hong, and Shue 2014; Flammer 2015, are notable exceptions incorporating natural or quasi-experiments). A frequent and credible variant of this approach employs ESG scores. These scores are issued by specialized rating agencies such as KLD/MSCI, Sustainalytics, or Asset4. For sub-dimensions (such as *emissions* or *human rights*), the well-known KLD/MSCI ESG assessments are available on a binary level only. To overcome this lack of variation of explanatory variables, we use data from Asset4, which offers continuous scores for various sub-dimensions and gained increasing interest in recent academic research (e.g. Stellner, Klein, and Zwergel 2015). We aggregate these scores to nine *area* scores covering the entire ESG universe. This novel information on the CSR of a company allows us to identify the genuine drivers of long-term stock performance.

Although ESG scores communicate investments in CSR activities, their valuation is rather ambiguous for the markets which contrasts with the efficient pricing hypothesis of such activities. Stock markets may not fully condition on CSR activities until they start to create additional positive cash flows several years after they have been initiated (Deng, Kang, and Low 2013) for at least two reasons: One, stock analysts are often myopically focused on quarterly published business figures, and two, the way in which socially responsible (SR) activities work are often rather opaque for the market. Therefore, we adopt a long-term perspective for analyzing the market reactions on CSR efforts.

We account for potential influential factors such as industry effects (see Khan, Serafeim, and Yoon 2016) and analyze the cross-section of stocks in a panel analysis, separately for each of five consecutive 12-months periods. To identify the particular types of intangibles, which are related with the market inefficiencies, we regress the abnormal returns for each time period on the area scores. We find that particularly *emission reduction* and *resource reduction* dominate the results in the environment dimension, while *workforce* and *society* stand out in the social dimension. Furthermore, we identify modest positive long-term abnormal returns predicted by high *vision & strategy* scores.

We find that the main economic channel determining the relationship between CSR and stock returns are additional cash flows, which we identify by earnings surprises. To overcome potential endogeneity concerns by analyzing CSR, earning surprises, and abnormal returns simultaneously, we implement a 2SLS approach and find that strong CSR predicts future earnings surprises. This provides clear evidence that the abnormal returns can partly be attributed to additional positive cash flows which are related to the respective CSR efforts. Additionally, we estimate the influence of a demand shift during our sample period towards strong CSR stocks and find that additional demand can account for an annual appreciation of 0.07–0.54%.

To test the economic significance of these results, we invest long in a portfolio of stocks with strong E, S, G, and certain area scores and short in a portfolio with weak E, S, G, and

certain area scores. Thereby, the scores are updated annually for each firm. In such an update, we consider the observation of a high (low) score as an event and analyze the abnormal returns during the subsequent 60 months. We again split the event window into five consecutive 12-months periods and calculate abnormal returns using both a characteristic-based benchmark portfolios approach and a calendar-time approach. In particular, we control for firm developments by risk factors such as size, book-to-market, and momentum throughout the entire holding period. We find that the E, S, and some area-ranked portfolios show statistically and economically significant abnormal returns in the period of 24–48 months after the portfolio formations.

Regarding the *financial performance* in the context of CSR, there are two notions to be distinguished. The first is the corporate financial performance (CFP) in the sense of the ability to sustainably generate cash flows, the second is the *stock market performance* of a company, i.e. the achievable (abnormal) returns by holding the shares of a company. While, unfortunately, some literature surveys do not clearly refer to these distinct concepts, in this article both notions essentially need to be differentiated. However, for both interpretations, there exists a huge number of published articles that analyze the relation between CSR and the CFP or the stock market performance, respectively. As our intention is to look for long-term abnormal stock returns, this study belongs to the second strand. Yet, we analyze the economic channels, thereby also tackling the CSR–CFP relation, and thus we build on both strands of research. A broad overview over the findings of the literature on stock market performance is given by Friede, Busch, and Bassen (2015) and von Wallis and Klein (2015). Overall, the findings from the pile of studies cited within these references are mixed and it appears that there is no clear evidence that investing in companies of investment funds with a high level of E, S, or G activity or awareness would yield to positive abnormal returns. Additionally, Halbritter and Dorfleitner (2015) find that some of the evidence that has been found in previous studies has vanished over time. In a recent article, Chen, Hung, and Lee (2017) indicate that firms with a high CSR value have a higher capacity to reduce their coordination costs with regard to shareholders and are more efficient in solving externalities and thus can create abnormal stock returns.

As we also partly argue with the CFP that is induced by CSR below, we are faced with a broad field of established hypotheses regarding the relationship between these two concepts. The first widespread argument is the *slack hypothesis*, which presumes a positive impact of financial performance on CSR. High financial performance yields financial slack that enables firms to invest in SR activities in the first place (Hong, Kubik, and Scheinkman 2012). In contrast, the *managerial opportunism hypothesis* refers to the agency problem of self-serving managers (Posner and Schmidt 1992). This leads to short-term maximization of private gains in prosperous times and to the placating of stakeholder disappointment in times of weak financial performance by increasing their welfare through the implementation of, for instance, social programs. While these two hypotheses consider the influence of financial performance on CSR, other papers argue in favor of a reverse relation. Firms can improve their reputation by catering to the implicit claims of major stakeholders (Cornell and Shapiro 1987) which results in a positive impact on financial performance. This hypothesis is called *good management hypothesis*. In contrast, the *trade-off hypothesis* stresses agency costs induced by CSR which yield a negative impact of CSR on financial performance (Aupperle, Carroll, and Hatfield 1985).

In a nutshell, we obviate endogeneity concerns raised by the *slack hypothesis* and the *managerial opportunism hypothesis* by attaching a long-term focus to our empirical framework and we provide clear evidence in favor of the *good management hypothesis*. However, due to market inefficiencies, stock prices do not reflect the increased financial performance induced by certain CSR activities immediately.

Continuing with the paper, in Section 2 we discuss possible economic channels of the materialization of abnormal returns by CSR in more detail and formalize two overarching hypotheses. In Section 3 we introduce our CSR data and derive our results on an ESG dimension level. We continue with Section 4, in which we use 2SLS regressions to identify cash flows as a second-order driver for the observed abnormal returns and separately estimate the magnitude of an additional demand channel. In Section 5 we identify exactly the ESG areas that are the first-order drivers of the abnormal returns and investigate the returns an investor can achieve by utilizing our findings, and in Section 6 we conclude our paper.

2. Theoretical background

Different CSR activities address different stakeholders and work through different economic channels. Orlitzky (2008) summarizes thirty years of research on the relationship of CSR activities and corporate financial performance. From this body of literature he identifies six CSR-related factors as the conceptual link being able to drive financial performance, i.e. to generate additional cash flows, namely an improvement of reputation and efficiency, higher costs to competing companies, a reduction of corporate risk, an increase in revenues, and the capability to attract a more productive workforce.

The *good management* and *trade-off hypotheses* conjecture an influence of CSR on corporate financial performance. When considering how CSR information is processed by the markets into stock prices, we are faced with three scenarios. One, a CSR investment has a positive or negative NPV which is priced efficiently by the markets. Two, a CSR investment has a positive or negative NPV which, however, is not priced efficiently by the markets. Three, a CSR investment has zero NPV and thus does not affect market prices.

Scenarios one and three suggest that markets are efficient, i.e. they incorporate the disclosed information in the stock prices immediately.[1] In this paper, we focus on scenario two in order to identify exactly those CSR activities that are related to market inefficiencies. We build our premise on several empirical studies which provide indications that stock markets are not fully efficient with respect to certain aspects of CSR and thus require some time to fully process the corresponding information due to its intangible nature (Edmans 2011; Edmans, Li, and Zhang 2014). To achieve an abnormal return in this context it is necessary that the activities have an opaque economic nature, implying that their future financial benefits are not immediately measurable. We refer to this rationalization as the *additional cash flow explanation* for abnormal returns.

To identify additional cash flows as a channel for abnormal returns we use earnings surprises, which are derived from analyst forecasts and known to be optimistically biased (see e.g. Gu and Wu 2003). Nevertheless, analyst forecasts include well-known information, which is generally priced efficiently. Hence, positive earnings surprises indicate the materialization of unexpected additional (positive) cash flows. The following hypothesis formalizes this rationalization:

H1: (Additional cash flows) Strong CSR stocks have positive long-term abnormal returns that are due to unexpected additional positive cash flows.

Next to changes on the supply side, the second economic channel for abnormal returns is a shift in demand to strong CSR stocks. Possible reasons for an additional demand are changes in the stock selection requirements of mutual funds due to severer screening policies with respect to ESG criteria (i.e. index effects) as well as rational and behavioral motives of investors. In particular, rational motives comprise effects such as the reduction of idiosyncratic risk (Derwall and Verwijmeren 2007), whereas behavioral motives comprise effects such as warm glow (Barnea and Rubin 2010). As the amount of high ESG stocks is limited on the markets, demand is concentrated within a limited number of stocks which then appreciate (Galema, Plantinga, and Scholtens 2008; Luo et al. 2015). We refer to this rationalization as the *additional demand explanation*. Since a shift in demand can persist in the mid run, it can lead to observed market inefficiencies when measured by abnormal returns (Stapleton and Subrahmanyam 1977). However, at some point in time a new equilibrium will be reached, causing the observed market inefficiencies due to additional demand to vanish. This rationalization yields our second hypothesis:

H2: (Additional demand) Strong CSR stocks have positive long-term abnormal returns that are due to additional demand.

From a corporate finance perspective, CSR affects costs of capital (Cheng, Ioannou, and Serafeim 2011; Ghoul et al. 2011; Goss and Roberts 2011; Klock, Mansi, and Maxwell 2005; Hoepner et al. 2016). Most studies find that strong CSR lowers both the systematic and the idiosyncratic risk of a company (Orlitzky 2008), resulting in lower debt and equity costs of capital. Since lower debt costs imply higher cash flows to equity, a change in debt costs is consistent with the additional cash flow hypothesis. Considering equity costs, there is ample evidence that markets are able to reflect the CSR-related risk information efficiently (Derwall and Verwijmeren 2007; Dhaliwal et al. 2011). Therefore, equity costs of capital are not within the scope of our paper. Lastly, lower costs of capital of high ESG stocks can result from an increased demand for these stocks. Given fixed future cash flows, an appreciation in stock prices yields lower costs of capital. This effect is directly accounted for in the additional demand hypothesis in our framework.

3. Methodology and basic results

In order to provide insights into the long-term effects of CSR on stock returns, our empirical framework is based on the computation of buy-and-hold abnormal returns (BHARs).

3.1. Methodology

For our basic results, we follow recent studies who recognize the heterogeneity of different CSR dimensions (see e.g. Khan, Serafeim, and Yoon 2016) and identify CSR by its classical dimensions, which are environment (E), social (S), and corporate government (G). We compute BHARs akin to Daniel et al. (1997) which control for the risk factors along size, book-to-market ratio, and momentum. As a characteristic-based approach it matches a stock along the mentioned risk factor to a benchmark portfolio that contains

stocks with similar characteristics.[2] At each calendar month, we use 64 ($= 4 \times 4 \times 4$) benchmark portfolios and calculate the abnormal return of a stock as being the difference between the stock return and its matching benchmark portfolio return. In our long-term analysis, for each stock i which is purchased in a certain month τ_0, we compute the BHAR during each of the five subsequent 12-months periods following τ_0. Formally, let $[\tau + 1, \tau + 12]$ be such a subsequent 12-months period with $\tau \in \{\tau_0, \tau_0 + 12, \tau_0 + 24, \tau_0 + 36, \tau_0 + 48\}$. We compute the BHAR of stock i in the period $[\tau + 1, \tau + 12]$ as

$$\mathrm{BHAR}^i(\tau_0,\tau) = \prod_{t=\tau+1}^{\tau+12} (1 + r_t^i) - \prod_{t=\tau+1}^{\tau+12} (1 + r_t^{P^i(\tau_0)}),$$

where r_t^i is the return of stock i in month t, $P^i(\tau_0)$ is the stock's value-weighted matching passive benchmark portfolio that was formed in τ_0, and $r_t^{P^i(\tau_0)}$ its return in month t. Notice that according to the criticism of Lyon, Barber, and Tsai (1999) on benchmark portfolios that are rebalanced periodically and therefore yield inflated long-horizon returns, our benchmark portfolio $P^i(\tau_0)$ is a strictly passive buy-and-hold portfolio over the entire holding period.[3] For each stock i, we use the months at whose beginning the stock's ESG score is updated (which happens once per year) as starting points τ_0. We then compute the BHAR for the sequence of the five consecutive 12-months periods $[\tau + 1, \tau + 12]$.

A major concern in studies investigating the CSR–abnormal-returns-relationship is reserve causality. It is hard to distinguish cause from effect when considering contemporary CSR and abnormal returns. In our long-term perspective this endogeneity problem does not occur since we analyze the impact of CSR in month τ_0 on abnormal returns in subsequent months. Profits in month τ_0 may influence CSR in τ_0. Nevertheless, since these profits are already priced in τ_0 they do not affect stock returns in subsequent months.

3.2. Data

Next to financial data, which we obtain from Thomson Reuters Datastream, comprehensive CSR data is essential for our analysis. Various specialized rating agencies provide either ESG scores or data on strengths and weaknesses with respect to CSR. Our endeavor will be to identify different areas of CSR as explanatory variables of future abnormal stock returns. Therefore, we require ESG data with an appropriate level of depth and variability: The depth ensures the coverage of the manifold fields of CSR activities. The variability of the specific metrics in each CSR dimension facilitates us to attain to statistical inference. The resource that provides us with such data in necessary depth and variability is Asset4, a Swiss sustainability rating agency. Asset4 publishes ratings for an international sample of over 5,000 companies beginning in 2002. To provide the variability of the ratings among the companies, each company rating is based upon more than 500 individual CSR metrics covering numerous aspects of sustainable performance. These metrics are aggregated in several steps to nine area scores (such as *emission reduction* and *workforce*) and to the three classical dimension E, S, and G scores. Each score ranges from 0 (worst) to 100 (best) and is standardized to a cross-sectional mean of 50.

The CSR metrics of Asset4 have a reputation for being among the most diligent and trustworthy sources of CSR data (Stellner, Klein, and Zwergel 2015). These CSR metrics are based on publicly available and traceable information such as websites, sustainability reports, media sources, SEC filings such as 10-K, and NGO reports. The rating approach of Asset4 to cross-check every entry by at least one additional analyst and by further analyses through statistical tools, guarantees a relatively high level of integrity of the data. Therefore, employing Asset4 CSR metrics mitigates weaknesses such as the lack of transparency in the KLD, FTSE4Good, and Dow Jones rating approaches (Chatterji and Levine 2006). This is due to the fact that Asset4 evaluates more than 500 individual data points. Every data point is linked to a single question regarding the fulfillment of a specific item according to environmental, social, and governance issues. The information gathered by the answers is aggregated in several stages, for instance to indicators and to the three dimensions of environment, social, and corporate governance. In particular, Asset4 applies considerable efforts to reflect the specific features of the three dimensions of ESG in their CSR metrics. These assessments are based on a variety of different indicators regarding emission and resource reduction and product innovation in the environment dimension, regarding workforce, society, and customers in the social dimension, and regarding board of directors, shareholder rights, and vision and strategy in the corporate governance dimension. The scores are updated on an annual basis. Asset4 also keeps a firm subsequent to bankruptcy, a merger, and other causes of delisting in its rating universe. Thus, the data set is free from survivorship bias. A detailed description on the CSR metrics of Asset4 is provided by Ioannou and Serafeim (2012) and Chatterji et al. (2016).

Our sample on CSR assessments comprises ratings of 1308 US and Canadian firms for at least one year during the period between 2002 and 2014. To alleviate a survivorship bias, we retain 257 firms in the sample which were delisted during the sample period. Ratings as well as all other data are obtained for the end of each month in our sample period. We filter our data with respect to two concerns to assure that our results are not driven by single outliers: First, we drop all stocks with a price of less than 1 USD in months τ_0 and we winsorize the 1 percent highest return observations. Second, we drop all strong/weak portfolios which contain less than 3 assets. After filtering our data, 1278 different firms remain in our sample. Table 1 reports upon descriptive statistics of our sample CSR data. While the average E and S score is approximately 10 points below the standardized world-wide average of 50, the average G score exceeds the average by approximately 20 points.

Former studies frequently use KLD/MSCI data to measure CSR. While KLD provides the longest time series of historical CSR data, this data lacks the granularity and variability needed in our study. Since KLD scores are binary variables on firm-specific strengths and weaknesses, and the bulk of smaller firms do not exhibit any strengths at all, the cross-

Table 1. Descriptive statistics on ESG scores.

	Mean	st. dev.	Skewness	Kurtosis	min	max
Environment score	41.67	30.67	0.55	1.69	8.30	97.28
Social score	44.63	28.46	0.34	1.78	3.57	98.88
Governance score	72.98	16.98	−1.23	4.73	1.43	98.78

Note: This table reports upon cross-sectional statistics of the three dimension scores.

sectional variation in these scores is rather limited. Therefore, the studies analyzed by Orlitzky, Schmidt, and Rynes (2003) and Margolis, Elfenbein, and Walsh (2009) may well under-identify an existing CSR–abnormal-returns-relation.

3.3. Cross-section of abnormal returns

To commence the empirical analysis, we analyze the entire cross-section of stocks and investigate the impact of ESG scores on abnormal returns while controlling for well-established determinants of stock returns. First, we control for the *debt ratio* as a main determinant of stock returns (see Bhandari 1988). Next, since several studies (Schwert 1983; Fama and French 1992) document the predictive power of firm size for returns, we control for firm size, calculated as the log value of *total assets*. Moreover, Desai and Jain (1997) document that stock splits, which lead to lower stock prices of the respective companies, yield positive long-run returns. Hence, we also include the *stock price* to our set of controls. Additionally, we include *profitability* and *investment*, as they represent essential components in the current five-factor model of Fama and French (2015).[4] Next, we add the *dividend-to-price ratio* as another determinant of stock returns (Fama and French 1988). Last, although the BHARs are already computed by matching stocks with a similar book-to-market value, we add the *book-to-market value* to our controls, as it might have additional predictive power as a factor (Daniel and Titman 1997; Pontiff and Schall 1998). We obtain annualized data for all controls from Datastream and Worldscope. Furthermore, since it is highly likely that the returns are affected by the respective industry and inhibit time-fixed patterns, we add industry and year fixed effects. Table 2 summarizes our set of controls and provides detailed definitions, while Table 3 provides descriptives statistics of the controls grouped by ESG score quintiles. Note that the remainder of the control variables is introduced in Section 4.1. The variation of the controls over the quintile groups emphasizes the necessity to control for these variables in the regressions.

Table 2. List of controls.

Variable	Description	Used for	
		BHAR	ES
debt ratio	Total debt (WC03255) divided by total assets (WC02999)	x	x
total assets	Natural logarithm of total assets (WC02999)	x	x
stock price	Natural logarithm of stock price (P)	x	x
profitability	Operating income (WC01250) divided by common shareholders' equity (WC03501)	x	x
investment	Total assets 1 year annual growth (WC08621)	x	x
dividend-to-price ratio	Cash dividend paid total (WC04551) divided by market value of equity (MV)	x	x
book-to-market value	Common shareholders' equity (WC03501) divided by market value of equity (MV)	x	x
year dummies	Dummy variable for each year	x	x
industry dummies	Industry by ICB classification model (INDM2)	x	x
number of analysts	Total number of estimators covering the company for the fiscal period		x
earnings volatility	Standard deviation of quarterly earnings per share during the past 4 years		x
earnings persistence	First-order autocorrelation coefficient of quarterly earnings per share during the past 4 years		x
reporting lag	Number of days from the quarter-end until the earnings announcement date		x
share turnover	Turnover by volume (VO)		x

Notes: This table contains detailed definitions of the control variables used in the cross-sectional regression models (Models A, A-2SLS, B-2SLS). In model A, we apply the controls marked in Column BHAR. in models A-2SLS and B-2SLS, we apply the controls marked in Column ES in the first stage and the controls marked in Column BHAR in the second stage. Datastream/Worldscope mnemonics are in parentheses where available.

Table 3. Descriptive statistics of controls.

ESG Score Quintile	1 (weak)	2	3	4	5 (strong)
debt ratio	0.2697	0.2560	0.2437	0.2452	0.2460
total assets	15.4084	15.3378	15.6225	16.2430	16.7793
stock price	3.3513	3.2964	3.3592	3.4015	3.5009
profitability	0.1886	0.1676	0.2706	0.4162	0.3363
investment	14.1723	12.3918	12.4457	8.5943	7.2703
dividend-to-price ratio	0.0298	0.0214	0.0185	0.0211	0.0240
book-to-market	0.7382	0.6006	0.5315	0.5800	0.5054

Notes: This table reports on descriptive statistics of the control variables grouped by ESG score quintiles. The ESG score is the average score of the three dimension scores.

In general, we set up the *cross-sectional regression analysis with dimension scores* model (Model A) which includes the BHARs as endogenous variables and the respective dimension score, the above controls including industry and year dummies as exogenous variables in a pooled OLS regression with firm-clustered standard errors.[5] In particular, we specify $3 \times 5 = 15$ variants of Model A, one for each of the three CSR dimensions and each of the five different 12-months holding periods starting in month t with $t \in \{1, 13, 25, 37, 49\}$ in event time. Table 4 summarizes the results. We report upon the coefficients of the respective CSR dimension, their t-statistics, and model statistics. All model specifications are significant. The results are based on 8539 firm-year observations for the holding period starting in the month after the ESG score update. The number of firm-year observations decreases with holding periods starting in the more distant future.

Table 4. Summary of cross-section analyses of CSR impact on BHAR on dimension level (Model A).

	1–12 months	13–24 months	25–36 months	37–48 months	49–60 months
	BHAR	BHAR	BHAR	BHAR	BHAR
			Environment (E) score		
E	0.0172	0.0190	0.0244*	0.0438***	0.0297*
	(1.31)	(1.34)	(1.88)	(3.08)	(1.83)
Obs.	8539	7987	7166	6314	5460
F-stat	4.760***	3.903***	3.102***	3.486***	3.814***
R^2	0.038	0.022	0.017	0.017	0.018
			Social (S) score		
S	0.0281*	0.0267*	0.0136	0.0234	0.0337**
	(1.90)	(1.76)	(0.96)	(1.54)	(2.02)
Obs.	8539	7987	7166	6314	5460
F-stat	4.670***	3.888***	3.108***	3.342***	4.012***
R^2	0.039	0.022	0.017	0.016	0.018
			Governance (G) score		
G	−0.0186	−0.0128	−0.0150	0.0044	0.0491*
	(−0.77)	(−0.54)	(−0.63)	(0.18)	(1.79)
Obs.	8539	7987	7166	6314	5460
F-stat	4.817***	3.935***	3.158***	3.358***	4.016***
R^2	0.038	0.022	0.017	0.016	0.018

Notes: This table contains the coefficients of the CSR dimensions in pooled OLS regressions with clustered standard errors with the BHAR as the dependent variable and the respective dimension of CSR, controls as described in Table 2, and both industry and year fixed effects as independent variables. Furthermore, this table reports upon the number of observations, the *F*-statistics of the models, and their R^2. ESG scores are interpreted as percentage values. t-statistics are given in parentheses.

*Significance at a 10% level.

**Significance at a 5% level.

***Significance at a 1% level.

In particular, *environment* activities yield positive mid-term and long-term abnormal returns amounting to a 3.03% ($= (0.0244 + 0.0438 + 0.0297) \times 0.3091$)) BHAR over years three to five for a one standard deviation change of the E score. *Social* activities also yield positive abnormal returns amounting to a 2.53% BHAR over years one, two, and five for a one standard deviation change of the S score. Note that all of the insignificant coefficients for both dimensions are strictly positive. *Governance* activities, however, yield (insignificant) negative abnormal returns in the years one to three and a 0.82% BHAR in year five for a one standard deviation change of the G score.[6]

4. Economic channels

The above results demonstrate that CSR activities are a predictor of future abnormal returns with respect to the environment and to the social dimensions. In our theoretical discussion, we have met two possible economic channels which are able to generate abnormal returns: The additional cash flow hypothesis and the additional demand hypothesis. To investigate the first, we identify unexpected additional cash flows by earnings surprises and analyze the influence of CSR on BHARs via earnings surprises. Given that earnings surprises are known to induce abnormal returns (Brown 1978; Watts 1978; Rendleman, Jones, and Latane 1982; Kane, Lee, and Marcus 1984), this facilitates us to quantify the effect of CSR-induced unexpected additional cash flows on abnormal returns. To quantify the scope of the effect of the second hypothesis, we estimate the influence of additional demand by employing both conservative and relaxed figures for the price elasticity of demand.

4.1. Additional cash flows

For the additional cash flow hypothesis, we use data from the Institutional Brokers' Estimate System (IBES) and follow Core, Guay, and Rusticus (2006), Edmans (2011), and Giroud and Mueller (2011) in calculating annual earnings surprises (ES) as

$$ES_{i,t} = \frac{Actual_{i,t} - MedianAnalyst_{i,t-8/12}}{Price_{i,t}},$$

where $Actual_{i,t}$ are the actual annual earnings per share, $MedianAnalyst_{i,t-8/12}$ the median analyst forecasts eight months before the earnings announcement date, and $Price_{i,t}$ the price (at the earnings announcement date) of stock *i*. In order to account for outliers, we winsorize our *ES* variable at the 1% and 99% level.

To simultaneously capture the effect of CSR on future earnings surprises and the influence of CSR and earnings surprises on BHARs we implement a simultaneous equations model, which we estimate via a 2SLS procedure (Model A-2SLS). At the first stage, we regress the earnings surprises on each of the three ESG score dimensions and a set of controls. As in Model A, we lag the explanatory variables by 12–60 months to pick up the long-run effect. At the second stage, the BHARs are regressed on the contemporary ES estimate and again on the past E, S, and G scores as well as the set of controls for BHARs (cf. Table 2). The controls for the earnings surprises comprise well-documented determinants, namely *number of analysts* (Brennan and Hughes 1991; Brown 1997),

earnings volatility (Donelson and Resutek 2015) and *earnings persistence* (Kothari, Lewellen, and Warner 2006), *reporting lag* (Chai and Tung 2002), and *share turnover* (Pfarrer, Pollock, and Rindova 2010). Additionally, we add all controls which are used for predicting the BHAR in the above regressions. Some of these are also known to be influential on ES (such as size and book-to-market), the remaining ones are added for consistency. As the sets of controls at both stages are thus not identical, we implicitly instrumentalize the endogenous variable ES by the five above-mentioned variables and therefore address potential endogeneity problems regarding ES.[7]

Generally, in this research setting different possible patterns can occur. Besides the outcome of insignificant coefficients, a significant influence of the respective E, S, or G score on the ES variable (first-order) and a significant positive relation of ES with the BHAR (second-order) at the same time indicates a ESG influence on the BHAR which is due to surprising additional cash flows. A significant coefficient of the E, S, or G score in the BHAR regression however, is an indication of additional demand rather than additional cash flows. Table 5 presents the results for each of the three ESG scores and all five time periods.

The findings of the 2SLS model can be summarized as follows. While the E and the S dimensions reveal a positive significant relation with the ES in all five periods, this translates into a positive impact on the BHAR in years two to five for both dimensions. There is also a modest indication of additional demand in the E dimension in year four, which accompanies the additional cash flows. To quantify the effect, we assert an additional BHAR over years two to five of 3.32% $(= (0.0048 \times 6.8153 + 0.0051 \times 7.5538 + 0.0042 \times 4.1739 + 0.005 \times 3.7382) \times 0.3091)$ for a one standard deviation change of the E score via the ES channel. For the social dimension we observe a 3.82% increase in BHAR over years two to five for a one-standard deviation change of the S score. There is no indication that the G score is related with ES in later years. However in year five a direct influence on BHAR can be observed, which could be due to additional demand. Therefore, the additional cash flow hypothesis (H1) is confirmed with respect to the E and the S score, while there is only limited evidence in favor of the additional demand hypothesis (H2).

4.2. Additional demand

With the above results already providing a modest indication for the *additional demand hypothesis*, we now examine the possible influence of an additional demand directly by estimating its impact on the stock prices of strong ESG stocks during our sample period.

According to US SIF (2014), the total net assets in mutual funds incorporating ESG factors increased from $162 billion at the end of 2002 to $1675 billion at the end of 2013. In order to estimate the shift in relative demand for strong ESG stocks, we need to put these figures into perspective with the total market capitalization of available strong ESG stocks in the respective years. To do so, we obtain all available portfolio holdings of mutual funds that are classified as incorporating ESG factors by the US SIF from the CRSP Survivor-Bias-Free U.S. Mutual Fund Database in 2002 (2013). The portfolios covered by this database invest in 816 (1934) different stocks which comprise a total market capitalization of $7171 billion ($20,451 billion) at the end of 2002 (2013). Therefore, the relative demand of strong ESG stocks increased from 2.26% in 2002 to 8.19% in 2013, i.e. there is a shift in demand of 5.93% towards strong ESG stocks. This demand

Table 5. Pooled 2SLS regressions (clustered standard errors) of ESG dimensions on ES and BHAR (Model A-2SLS).

	1–12 months		13–24 months		25–36 months		37–48 months		49–60 months	
	ES	BHAR	ES	BHAR	ES	BHAR	ES	BHAR	ES	BHAR
					Environment (E) score					
E	0.0052***	0.0433	0.0048**	− 0. 0170	0.0051***	− 0. 0130	0.0042**	0.0268*	0.0050**	0.0101
	(2.79)	(0.90)	(2.41)	(−0.70)	(2.59)	(−0.69)	(2.17)	(1.62)	(2.35)	(0.50)
ES		−4.472		6.8153*		7.5538***		4.1739**		3.7382*
		(−0.54)		(1.93)		(4.03)		(2.37)		(1.82)
Obs.	8539		7987		7166		6314		5460	
F-stat	9.131***		6.344***		6.126***		5.932***		5.522***	
R^2	0.085	.	0.060	.	0.067	.	0.065	0.027	0.064	0.056
					Social (S) score					
S	0.0069***	0.0565	0.0068***	−0.0277	0.0057***	− 0. 0237	0.0048**	0.0060	0.0047*	0.0161
	(3.58)	(1.02)	(3.56)	(−0.95)	(2.80)	(−1. 28)	(2.12)	(0.36)	(1.89)	(0.85)
ES		−3.7617		7.8289**		7.4061***		4.1874**		3.8568*
		(−0.48)		(2.36)		(4.12)		(2.49)		(1.92)
Obs.	8.539		7.987		7.166		6.314		5.460	
F-stat	9.383***		6.432***		6.192***		5.975***		5.503***	
R^2	0.086	.	0.061	.	0.067	.	0.065	0.025	0.064	0.051
					Governance (G) score					
G	0.0035	− 0. 0078	0.0032	− 0. 0341	− 0. 0023	0.0032	− 0. 0022	0.0136	− 0. 0019	0.0545*
	(1.34)	(−0.22)	(1.11)	(−1. 19)	(−0.76)	(0.10)	(−0.72)	(0.54)	(−0.55)	(1.90)
ES		−2.6072		6.0391*		7.6246***		4.1117**		3.5362*
		(−0.39)		(1.85)		(3.88)		(2.37)		(1.74)
Obs.	8539		7987		7166		6314		5460	
F-stat	9.228***		6.322***		6.131***		5.929***		5.380***	
R^2	0.084	.	0.059	.	0.066	.	0.065	0.030	0.063	0.065

Notes: This table contains the coefficients of the ESG scores and the earnings surprises of five 2SLS regressions (for the five time brackets) with the ES of the corresponding period as the dependent variable in the first step and BHAR in the second step. Besides the displayed variables several controls (cf. Table 2) enter the regression as independent variables. Furthermore, this table reports upon the number of observations, the *F*-statistics of the models, and their R^2. ESG scores are interpreted as percentage values. *t*-statistics are given in parentheses.

*Significance at a 10% level.

**Significance at a 5% level.

***Significance at a 1% level.

change affects stock prices by

$$\Delta P = 0.0593/\varepsilon,$$

where ε denotes the absolute price elasticity of demand. While a perfect market presumes infinite elasticity, Wurgler and Zhuravskaya (2002) estimate $\varepsilon = 8$ using demand shocks induced by S&P 500 constituents changes. A very conservative measure is a unit elasticity (Edmans 2011). Using a range of 1–8 for the elasticity, we estimate the abnormal return induced by the additional demand between 2002 and 2013 to be in the range of 0.74–5.93%, or 0.07–0.54% per year. Given the magnitude of these figures, we conclude that next to unexpected additional cash flows, also additional demand can account for the observed BHARs to a certain extent.

5. Digging deeper: the drivers of profitable ESG investing and returns of a zero-investment strategy

While the E, S, and G scores reflect the classical CSR dimensions, each dimension comprises activities in several areas. The activities in each area address different stakeholders and yield differently structured streams of cash flows. In order to identify exactly the areas that drive the previous results, we extend Model A-2SLS, replacing each of the broader E, S, and G exogenous variables by a set of more granular area assessments. Moreover, we confirm the economic significance of the results by considering long-term portfolio strategies.

5.1. Data and methodology

In particular, we follow the area classification of Asset4. This classification consists of three areas for each of the three dimensions, respectively, in particular *resource reduction*, *emission reduction*, and *product innovation* in the environment dimension, *workforce*, *society*, and *customer* in the social dimension, and *board of directors*, *shareholder rights*, and *vision & strategy* in the governance dimension. We provide a detailed description of the $3 \times 3 = 9$ areas in Table 6. Moreover, Table 7 reports descriptive statistics on each of the area scores. While the average area scores within the environment and social dimension are similar to the average dimension scores (cf. Table 1), the areas within the governance dimension exhibit more heterogeneous average scores. In particular, the average *vision & strategy* score of 39.83 is substantially below the average governance dimension score of 72.98.

With these more granular data, we set up the *cross-sectional regression analysis with area scores* model (Model B-2SLS), which is analogous to Model A-2SLS in Table 5, but with each of the three respective areas of each dimensions as explanatory variables instead of the whole dimension itself. For each ESG dimension, we now specify three variants of Model B-2SLS for each of the five consecutive 12-months holding periods, i.e. we are running $3 \times 3 \times 5 = 45$ regressions. Tables 8–10 present the results for each dimension, respectively.

5.2. Results and discussion

The results of the 2SLS analysis on the area level provide additional insights. First, in the environment dimension (Table 8) both reduction areas can produce abnormal returns via the ES channel. We identify a significant positive long-term abnormal return of the

Table 6. Area descriptions.

	Environment dimension
Resource reduction (*ENRR*)	Commitment and effectiveness towards achieving an efficient use of natural resources in the production process. Capacity to reduce the use of materials, energy or water, and to find more eco-efficient solutions by improving supply chain management.
Emission reduction (*ENER*)	Commitment and effectiveness towards reducing environmental emission in the production and operational processes. Capacity to reduce air emissions, waste, hazardous waste, water discharges, spills or its impacts on biodiversity and to partner with environmental organizations to reduce the environmental impact of the company in the local or broader community.
Product innovation (*ENPI*)	Commitment and effectiveness towards supporting the research and development of eco-efficient products or services. Capacity to reduce the environmental costs and burdens for its customers, and thereby creating new market opportunities through new environmental technologies and processes or eco-designed, dematerialized products with extended durability.
	Social dimension
Workforce $\frac{1}{4}(SODO + SOEQ + SOHS + SOTD)$	Commitment and effectiveness towards providing high-quality employment benefits and job conditions, providing a healthy and safe workplace, providing training and development (education) for its workforce, and maintaining diversity and equal opportunities in its workforce. Capacity to increase its workforce loyalty and productivity by distributing rewarding and fair employment benefits, by focusing on long-term employment growth and stability by promoting from within, avoiding lay-offs and maintaining relations with trade unions, by integrating into its day-to-day operations a concern for the physical and mental health, well-being and stress level of all employes, and by promoting an effective life-work balance, a family friendly environment and equal opportunities regardless of gender, age, ethnicity, religion or sexual orientation. Capacity to increase its intellectual capital, workforce loyalty and productivity by developing the workforce's skills, competences, employability and careers in an entrepreneurial environment.
Society $\frac{1}{2}(SOCO + SOHR)$	Commitment and effectiveness towards respecting the fundamental human rights conventions and maintaining the company's reputation within the general community (local, national and global). Capacity to maintain its license to operate by guaranteeing the freedom of association and excluding child, forced or compulsory labor, by being a good citizen (donations of cash, goods or staff time, etc.), protecting public health (avoidance of industrial accidents, etc.) and respecting business ethics (avoiding bribery and corruption, etc.).
Customer (*SOPR*)	Commitment and effectiveness towards creating value-added products and services upholding the customer's security. Capacity to maintain its license to operate by producing quality goods and services integrating the customer's health and safety, and preserving its integrity and privacy also through accurate product information and labeling.
	Governance dimension
Board of directors $\frac{1}{3}(CGBF + CGBS + CGCP)$	Commitment and effectiveness towards following best practice corporate governance principles related to a well balanced membership of the board, following best practice corporate governance principles related to competitive and proportionate management compensation, and following best practice corporate governance principles related to board activities and functions. Capacity to ensure a critical exchange of ideas and an independent decision-making process through an experienced, diverse and independent board. Capacity to attract and retain executives and board members with the necessary skills by linking their compensation to individual or company-wide financial or extra-financial targets. Capacity to have an effective board by setting up the essential board committees with allocated tasks and responsibilities.
Shareholder rights (*CGCR*)	Commitment and effectiveness towards following best practice corporate governance principles related to a shareholder policy and equal treatment of shareholders. Capacity to be attractive to minority shareholders by ensuring them equal rights and privileges and by limiting the use of anti-takeover devices.
Vision & strategy (*CGVS*)	Commitment and effectiveness towards the creation of an overarching vision and strategy integrating financial and extra-financial aspects. Capacity to convincingly show and communicate that it integrates the economic (financial), social and environmental dimensions into its day-to-day decision-making processes.

Notes: This table provides detailed descriptions of the 9 areas employed. The calculation of the area scores is displayed by the italic Datastream mnemonics. Source: Asset4 ESG Data Glossary.

Table 7. Descriptive statistics on ESG area scores.

	mean	st. dev.	skewness	kurtosis	min	max
			Environment dimension			
Resource reduction	44.57	32.07	0.51	1.62	6.81	97.36
Emission reduction	42.61	31.10	0.34	1.45	7.27	97.95
Product innovation	44.26	29.77	0.67	1.82	7.82	99.68
			Social dimension			
Workforce	47.29	22.02	0.27	2.06	4.47	97.62
Society	46.86	25.02	0.39	2.01	7.66	98.27
Customer	48.23	27.93	0.27	1.74	2.40	99.06
			Governance dimension			
Board of directors	75.29	10.47	−1.69	7.63	5.55	94.48
Shareholder rights	64.54	26.20	−0.40	1.87	1.01	99.13
Vision & strategy	39.83	30.19	0.73	1.92	8.40	98.64

Note: This table reports upon cross-sectional statistics of the nine area scores.

emission reduction as well as of the *resource reduction* score in the period of two to five years after an increase in the score. To quantify the effect, we assert an ES-driven BHAR over years two to five of 3.14% in the *resource reduction* area and 2.99% in the *emission reduction* area for a one standard deviation change of the respective score. However, we cannot completely disentangle the influences of the *emission reduction* score from influences of the *resource reduction* score, since the two scores are highly correlated with a Pearson correlation of 85.2%. Yet, our results provide clear evidence that the abnormal returns are mainly driven by the reduction areas. While the influence of *product innovation*, on the other hand, is statistically significant in the respective years, its economic significance (measured by the induced BHAR by a one standard deviation change) is substantially lower. In summary, particularly the internalization of externalities in the environment dimension appears to be rewarded by the markets in the long run.

The social dimension of CSR comprises three areas, with each relating to different stakeholders, namely *workforce*, *society*, and *customer*. The results of our 2SLS model are displayed in Table 9. The *workforce* area yields a significant positive effect for all years after the 1-to-12 months bracket, amounting to a 3.62% ES-induced BHAR over years two to four for a one standard deviation change of the workforce score. This finding is consistent with Edmans (2011), who reports similar evidence analyzing Fortune's '100 Best Companies (BC) to Work For in America' list. Indeed, after matching this list to our data, the BC list constituents exhibit an average workforce area score of 63.1, which is significantly higher (at a 1% level) than the respective average score of 44.3 of all non-listed firms. Comparing the term structure of the BHARs to those of Edmans (2011), we find evidence that the BC list membership is lagged compared to the workforce scores. In our sample, the positive BHARs still persist in the 49-to-60 months bracket while for the BC list they vanish after 48 months. Moreover, our results are consistent with Galema, Plantinga, and Scholtens (2008), who document a modest positive influence of employe relations on excess returns. In the *society* area, which is positively correlated with the workforce area, we find very similar results. However, the ES-induced BHAR over years two to five for a one standard deviation change of the society score amounts to only 3.08%, which indicates that the *workforce* area score has a higher economic significance. Although customers are a significant stakeholder for most companies (Luo and Bhattacharya 2006), the *customer* area score in the context of CSR measures primarily product responsibility

Table 8. Pooled 2SLS regressions (clustered standard errors) of the environment areas on ES and BHAR (Model B-2SLS).

	1–12 months		13–24 months		25–36 months		37–48 months		49–60 months	
	ES	BHAR	ES	BHAR	ES	BHAR	ES	BHAR	ES	BHAR
					Resource reduction (RR) score					
RR	0.0050***	0.0391	0.0042**	−0.0136	0.0051***	−0.0224	0.0037*	0.0213	0.0043**	0.0061
	(2.76)	(0.91)	(2.17)	(−0.62)	(2.59)	(−1.18)	(1.83)	(1.31)	(1.99)	(0.34)
ES		−3.9562		6.6038*		7.5723***		4.1886**		3.7564*
		(−0.50)		(1.91)		(4.06)		(2.39)		(1.84)
Obs.	8539		7987		7166		6314		5460	
F-stat	9.170***		6.371***		6.117***		5.922***		5.528***	
R^2	0.085	.	0.060	.	0.067	.	0.065	0.026	0.064	0.055
					Emission reduction (ER) score					
ER	0.0064***	0.0406	0.0046**	−0.0176	0.0045**	−0.0086	0.0036*	0.0093	0.0043**	0.0076
	(3.35)	(0.72)	(2.26)	(−0.73)	(2.28)	(−0.46)	(1.83)	(0.58)	(2.11)	(0.39)
ES		−4.2182		6.7484*		7.5701***		4.1505**		3.7519*
		(−0.51)		(1.94)		(4.00)		(2.39)		(1.82)
Obs.	8539		7987		7166		6314		5460	
F-stat	9.132***		6.331***		6.102***		5.922***		5.529*** *	
R^2	0.086	.	0.060	.	0.067	.	0.065	0.027	0.064	0.056
					Product innovation (PI) score					
PI	0.0025	0.0229	0.0035*	−0.0141	0.0033*	−0.0053	0.0033*	0.0255	0.0034*	0.0003
	(1.38)	(0.92)	(1.93)	(−0.72)	(1.94)	(−0.31)	(1.83)	(1.64)	(1.86)	(0.02)
ES		−3.4973		6.1966*		7.5433***		4.1239**		3.6923*
		(−0.48)		(1.80)		(3.94)		(2.35)		(1.83)
Obs.	8539		7987		7166		6314		5460	
F-stat	9.067***		6.308***		6.175***		5.928***		5.402***	
R^2	0.084	.	0.060	.	0.067	.	0.065	0.029	0.064	0.058

Notes: This table contains the coefficients of the E area scores and the earnings surprises of five 2SLS regressions (for the five time brackets) with the ES of the corresponding period as the dependent variable in the first step and BHAR in the second step. Besides the displayed variables several controls (cf. Table 2) enter the regression as independent variables. Furthermore, this table reports upon the number of observations, the *F*-statistics of the models, and their R^2. ESG scores are interpreted as percentage values. *t*-statistics are given in parentheses.

*Significance at a 10% level.

**Significance at a 5% level.

***Significance at a 1% level.

Table 9. Pooled 2SLS regressions (clustered standard errors) of the social areas on ES and BHAR (Model B-2SLS).

	1–12 months		13–24 months		25–36 months		37–48 months		49–60 months	
	ES	BHAR	ES	BHAR	ES	BHAR	ES	BHAR	ES	BHAR
					Workforce (WO) score					
WO	0.0091***	0.0667	0.0085***	−0.0480	0.0073***	−0.0400*	0.0055*	0.0173	0.0058**	0.0172
	(3.56)	(1.03)	(3.25)	(−1.36)	(2.69)	(−1.68)	(1.95)	(0.80)	(2.02)	(0.73)
ES		−3.0240		7.6303**		7.3329***		4.2893**		3.8994*
		(−0.42)		(2.40)		(4.17)		(2.53)		(1.96)
Obs.	8539		7987		7166		6314		5460	
F-stat	9.206***		6.365***		6.194***		5.980***		5.457***	
R^2	0.086	.	0.061	.	0.067	.	0.065	0.020	0.064	0.049
					Society (SO) score					
SO	0.0050**	0.0381	0.0050**	−0.0062	0.0058***	−0.0195	0.0051**	0.0025	0.0061**	−0.0057
	(2.39)	(0.85)	(2.45)	(−0.22)	(2.63)	(−0.98)	(2.11)	(0.13)	(2.21)	(−0.26)
ES		−3.7256		6.8314*		7.6061***		4.1679**		3.8753*
		(−0.48)		(1.95)		(4.05)		(2.43)		(1.91)
Obs.	8539		7987		7166		6314		5460	
F-stat	9.504***		6.426***		6.203***		5.945***		5.497***	
R^2	0.085	.	0.060	.	0.067	.	0.065	0.026	0.064	0.050
					Customer (CU) score					
CU	0.0037**	0.0172	0.0022	−0.0085	0.0011	0.0085	0.0027	−0.0155	0.0007	0.0085
	(2.28)	(0.62)	(1.37)	(−0.56)	(0.62)	(0.53)	(1.40)	(−1.06)	(0.33)	(0.56)
ES		−2.8369		6.0217*		7.7042***		3.8675**		3.7420*
		(−0.42)		(1.87)		(4.01)		(2.27)		(1.87)
Obs.	8539		7987		7166		6314		5460	
F-stat	9.130***		6.358***		6.181***		5.943***		5.469***	
R^2	0.085	.	0.059	.	0.066	.	0.065	0.042	0.063	0.056

Notes: This table contains the coefficients of the S area scores and the earnings surprises of five 2SLS regressions (for the five time brackets) with the ES of the corresponding period as the dependent variable in the first step and BHAR in the second step. Besides the displayed variables several controls (cf. Table 2) enter the regression as independent variables. Furthermore, this table reports upon the number of observations, the *F*-statistics of the models, and their R^2. ESG scores are interpreted as percentage values. *t*-statistics are given in parentheses.

*Significance at a 10% level.

**Significance at a 5% level.

***Significance at a 1% level.

Table 10. Pooled 2SLS regressions (clustered standard errors) of the governance areas on ES and BHAR (Model B-2SLS).

	1–12 months		13–24 months		25–36 months		37–48 months		49–60 months	
	ES	BHAR	ES	BHAR	ES	BHAR	ES	BHAR	ES	BHAR
					Board of directors (BD) score					
BD	− 0. 0039	− 0. 0765	− 0. 0006	− 0. 0341	− 0. 0055	0.0186	− 0. 0051	0.0035	− 0. 0048	0.0756*
	(−1. 01)	(−1. 52)	(−0.14)	(−0.97)	(−1. 22)	(0.44)	(−1. 13)	(0.10)	(−0.98)	(1.89)
ES		−2.4717		5.6824*		7.6049***		4.0875**		3.4969*
		(−0.38)		(1.76)		(3.88)		(2.37)		(1.74)
Obs.	8539		7987		7166		6314		5460	
F-stat	9.075***		6.293***		6.140***		5.933***		5.390***	
R^2	0.084	.	0.059	.	0.066	.	0.065	0.031	0.063	0.066
					Vision & strategy (VS) score					
VS	0.0043**	0.0246	0.0037*	− 0. 0190	0.0030	− 0. 0160	0.0027	0.0201	0.0045**	0.0115
	(2.27)	(0.68)	(1.96)	(−0.89)	(1.62)	(−0.98)	(1.34)	(1.24)	(2.11)	(0.59)
ES		−3.5247		6.3254*		7.5814***		4.1607**		3.7258*
		(−0.47)		(1.85)		(3.99)		(2.37)		(1.83)
Obs.	8539		7987		7166		6314		5460	
F-stat	9.337***		6.376***		6.138***		5.909***		5.525***	
R^2	0.085	.	0.060	.	0.066	.	0.065	0.027	0.064	0.057
					Shareholder rights (SR) score					
SR	0.0038**	0.0174	0.0021	− 0. 0038	− 0. 0012	0.0105	− 0. 0009	0.0067	− 0. 0013	0.0173
	(2.07)	(0.58)	(1.19)	(−0.22)	(−0.66)	(0.57)	(−0.41)	(0.43)	(−0.53)	(0.98)
ES		−2.6449		5.7103*		7.6545***		4.1391**		3.6913*
		(−0.40)		(1.76)		(3.91)		(2.40)		(1.85)
Obs.	8539		7987		7166		6314		5460	
F-stat	9.169***		6.342***		6.126***		5.908***		5.378***	
R^2	0.085	.	0.059	.	0.066	.	0.065	0.028	0.063	0.058

Notes: This table contains the coefficients of the G area scores and the earnings surprises of five 2SLS regressions (for the five time brackets) with the ES of the corresponding period as the dependent variable in the first step and BHAR in the second step. Besides the displayed variables several controls (cf. Table 2) enter the regression as independent variables. Furthermore, this table reports upon the number of observations, the *F*-statistics of the models, and their R^2. ESG scores are interpreted as percentage values. *t*-statistics are given in parentheses.

*Significance at a 10% level.

**Significance at a 5% level.

***Significance at a 1% level.

(cf. Table 6). Our results show that the activities regarding product responsibility (*customer*) have no significant impact on BHARs during all time brackets.

The governance dimension comprises three areas: *board of directors*, *shareholders*, and *vision & strategy*. The results are reported in Table 10 and extend the previous observation from Model A-2SLS that no ES-driven abnormal returns can be achieved in this dimension (cf. Table 5). As detailed in Table 6, the *board of directors* score comprises activities in three different categories, i.e. board structure, compensation policy, and board functions. Our results in this area are consistent with Bhagat and Bolton (2008), who provide evidence that many governance aspects are priced reasonably efficiently by the market. Only the *vision & strategy* (VS) area, which measures the efforts undertaken by management to integrate CSR activities into the day-to-day decision making process, affects ES in the 13-to-24 and the 49-to-60 months brackets with an ES-induced BHAR of 1.21% for a one-standard deviation change in the VS score.[8] Regarding *shareholder rights*, we do not expect any influence of the score on future BHARs due to the absence of informational asymmetries, as there are no informational frictions between market participants and the stakeholders affected by shareholder rights (since these two groups coincide). In line with our expectations and with Bhagat and Bolton (2008), we find no evidence that *shareholder rights* was a significant driver of positive BHARs.

5.3. Zero investment returns

In this subsection, we confirm the significance of our findings from an economic perspective by setting up tradable zero-investment strategies to exploit the findings. We construct five quintile portfolios for each E, S, and G score. Notice that the 1st quintile portfolio represents the *weak* portfolio and the 5th quintile portfolio represents the *strong* portfolio. Figure 1 displays the average cumulative abnormal returns for a five-year holding period for each of these portfolios. The weak portfolios exhibit negative abnormal returns, whereas the strong portfolios exhibit positive abnormal returns. Generally, Figure 1 suggests a positive correlation between CSR and long-term abnormal returns in each CSR dimension.

Next, we conduct a zero investment portfolio analysis for the three dimensions (E, S, and G) and the five areas with significant results in the cross-sectional analysis (emission reduction, resource reduction, workforce, society, and vision & strategy). We form eight zero investment portfolios which invest long in those stocks representing the highest dimension or area scores quintiles, respectively, and short in those stocks representing

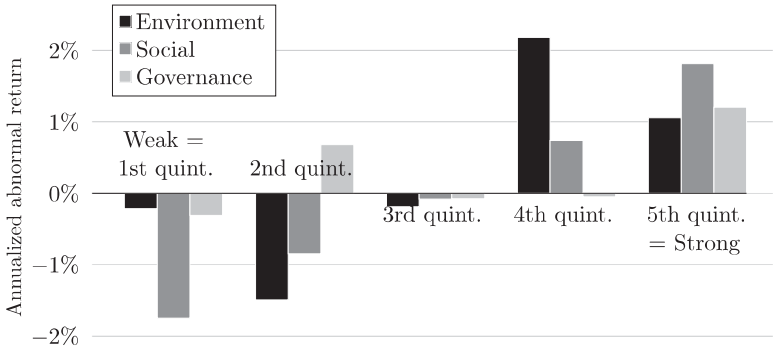

Figure 1. Average 5-year cumulative buy-and-hold abnormal returns on quintile portfolios formed on three sustainable scores: 01/2002–12/2014.

the lowest respective scores quintiles. Following the zero investment strategy, we construct equally weighted strong and weak portfolios for each month τ_0 and each of the dimension and area scores $D \in$ {E,S,G,ER,RR,WO,SO,VS}. The strong (weak) D-portfolio at month τ_0 includes the 20% best (worst) stocks regarding the respective score at the end of month τ_0. The BHAR of the zero investment strategy starting in month τ_0 with a 12-months holding period from $\tau + 1$ through $\tau + 12$ is

$$\mathrm{BHAR}^{D}(\tau_0,\tau) = \mathrm{BHAR}^{\mathrm{strong}_D}(\tau_0,\tau) - \mathrm{BHAR}^{\mathrm{weak}_D}(\tau_0,\tau),$$

where $\mathrm{BHAR}^{\mathrm{strong}_D}(\tau_0,\tau)$ and $\mathrm{BHAR}^{\mathrm{weak}_D}(\tau_0,\tau)$ are the BHARs of the equally weighted strong and weak portfolios of score $D \in$ {E,S,G,ER,RR,WO,SO,VS}. Finally, we compute the average zero investment BHAR over all starting months τ_0 in our sample period for each of the three dimension scores.

Panel (a) of Table 11 reports upon the results of the five consecutive 12-months holding periods. We test the abnormal returns for statistical inference using Newey-West heteroscedasticity-consistent standard errors. When inspecting the BHARs for the eight dimension and area scores strong-minus-weak (SMW) portfolios, we find positive abnormal returns of the applied trading strategies. Besides the earlier-reported effect for the workforce area (Edmans 2011), we document strong indications of superior long-term performance of zero investment strategies based on each dimension and most of the area scores. In particular, this findings show that ESG scores provide valuable information proxies for long-term trading strategies to beat the market portfolio significantly. The level of the success of the respective strategy depends on the selected area, respectively, dimension.

To assure that the above results are robust to the applied methodology of computing BHARs, we also compute calendar-time regression alphas with monthly data. Here, we follow the methodology of Peyer and Vermaelen (2009) and compute the alphas of the calendar-time regressions

$$r_{P(t,\tau)} = \alpha + \beta_1(Mkt_t - r_t^f) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 WML_t + \varepsilon_t.$$

In this equation, $P(t,\tau)$ denotes the equal-weighted portfolio that contains long (short) all stocks which were among the highest (lowest) 20% of stocks in the respective score at least once in the period from $t - \tau - 12$ through $t - \tau - 1$. To again consider five consecutive 12-months periods subsequent to a stock being in a strong/weak portfolio, we let $\tau \in$ {0, 12, 24, 36, 48}. Furthermore, Mkt_t denotes the return of the market portfolio in month t, SMB_t denotes the return of the small-minus-big portfolio in month t, HML_t denotes the return of the high-minus-low portfolio in month t, WML_t denotes the return of the winners-minus-losers portfolio in month t, and r_t^f denotes the risk-free rate in month t.[9]

Panel (b) of Table 11 reports upon the calendar-time alphas with the computed BHARs as endogenous variables.[10] Whilst this is a conservative approach, it may still shed some light on the results as characteristics such as size, book-to-market, and momentum can still have explanatory power, even after the returns have been controlled for covariances with these characteristics (Daniel and Titman 1997). While the statistical inference differs in several periods, the results from Panel (b) are generally consistent with the results from Panel (a).

Table 11. Abnormal returns of strong-minus-weak portfolios.

Panel (a): Pure BHARs

	1–12 months	13–24 months	25–36 months	37–48 months	49–60 months
	BHAR	BHAR	BHAR	BHAR	BHAR
E score	0.0088	−0.0001	0.0380**	0.0223	−0.0054
	(0.40)	(−0.01)	(1.97)	(1.35)	(−0.31)
S score	0.0107	0.0311	0.0669***	0.0528***	0.0165
	(0.40)	(1.61)	(4.39)	(3.37)	(0.97)
G score	−0.0041	0.0378*	0.0315	0.0145	−0.0039
	(−0.19)	(1.81)	(1.52)	(0.67)	(−0.17)
Resource Reduction	0.0098	0.0134	0.0155	0.0217	0.0044
	(0.55)	(0.90)	(0.95)	(1.47)	(0.29)
Emission Reduction	0.0044	0.0245	0.0263*	0.0326*	−0.0155
	(0.25)	(1.39)	(1.71)	(1.90)	(−0.91)
Workforce	0.0172	0.0370**	0.0673***	0.0788***	0.0106
	(0.65)	(2.41)	(4.14)	(4.62)	(0.56)
Society	0.0186	0.0378*	0.0540***	0.0241	0.0277
	(0.95)	(1.95)	(3.25)	(1.38)	(1.35)
Vision & Strategy	0.0197	0.0299*	0.0480**	0.0204	−0.0050
	(1.21)	(1.65)	(2.52)	(1.12)	(−0.36)

Panel (b): Calendar-time regression on BHARs

	alpha	alpha	alpha	alpha	alpha
E score	0.0154	−0.0024	0.0249*	0.0145	0.0366**
	(1.00)	(−0.13)	(1.65)	(0.88)	(2.05)
S score	0.0387***	0.0200	0.0285	0.0443***	0.0315**
	(2.58)	(1.29)	(1.55)	(2.81)	(2.03)
G score	0.0146	0.0321*	0.0242*	0.0255*	0.0262*
	(1.30)	(1.89)	(1.67)	(1.68)	(1.68)
Resource Reduction	0.0191	0.0104	0.0141	0.0170	0.0177
	(1.40)	(0.78)	(1.07)	(1.31)	(1.26)
Emission Reduction	0.0064	0.0144	0.0095	0.0102	0.0139
	(0.45)	(1.04)	(0.64)	(0.66)	(0.87)
Workforce	0.0252	0.0005	0.0246	0.0363**	0.0334**
	(1.57)	(0.03)	(1.36)	(2.16)	(2.01)
Society	0.0152	0.0068	0.0237	0.0400**	0.0367**
	(1.06)	(0.32)	(1.16)	(2.04)	(2.02)
Vision & Strategy	0.0138	0.0124	0.0181	0.0182	0.0143
	(1.18)	(1.02)	(1.45)	(1.33)	(1.06)

Notes: The table reports upon the abnormal returns of portfolios, consisting of long/short positions in stocks that are among the best/worst 20% with respect to the environment score, the social score, and the governance score, respectively. The sample period consists of all ESG-rated North American firms from 2002–2014. The abnormal returns are measured in two different ways: Panel (a) uses BHARs as in Daniel et al. (1997) and Panel (b) uses calendar-time regressions on the BHARs from Panel (a). The alphas in panel (b) are annualized by multiplying the monthly alpha by 12. In parentheses are *t*-statistics computed using newey–West heteroscedasticity-consistent standard errors.

*Significance at a 10% level.

**Significance at a 5% level.

***Significance at a 1% level.

6. Conclusions

In this paper, we analyze long-term abnormal returns of different dimensions of CSR. We are able to document significant mid- and long-term predictive power for the environment and social dimensions while controlling for various explanatory factors. These findings are confirmed in zero-investment portfolio analysis. We demonstrate that the additional value is not only generated by additional demand to high-CSR stocks during our sample period, but that firms with high CSR are indeed able to increase their operational performance by generating unanticipated cash flows.

The findings of this paper have profound implications for both asset management and corporate finance. In asset management, fund managers can expect to generate abnormal returns by investing in firms that exhibit a high CSR in the respective areas and by holding the stocks for a longer period. Yet, some caution is advised since a part of the abnormal returns identified in this paper is attributed to additional demand which cannot be extrapolated to the future. The areas which yield an ES-driven long-term abnormal return are *emission reduction* and *resource reduction* in the environment dimension, *workforce* and *society* in the social dimension, and to a less pronounced extent the *vision & strategy* area in the governance dimension.

From a corporate finance perspective, firms can increase their fundamental value by investing in the CSR activities identified in this paper. However, markets appear to condition on these investments into intangibles only after they start to generate tangible assets in the form of unexpected additional cash flows. Given the efforts firms spend on disclosing their ESG activities, it seems unlikely that the non-incorporation of CSR-intangibles into market prices is due to information asymmetries, but it also appears to be a rather inherent attribute of the markets. In this sense, this paper contributes to the stream of papers written on market short-termism.

Considering the causal relation between CSR and abnormal returns, we rule out reverse causality between positive BHARs and CSR several years earlier as an alternative explanation for the observed relations for two reasons. First, it would be necessary that management privately expects additional cash flows emerging in future years without communicating and disclosing this expectation to the markets. Given the short-termism of both markets and management performance remuneration, this seems highly unlikely. Second, even assuming that management has both private information about additional future cash flows and incentives not to disclose it, management would need to be able to convert this knowledge into current funds that are required to implement the CSR activities. Both possible avenues – issuing currently underpriced shares (given the knowledge about future superior stock returns) or raising debt at currently overestimated costs of capital – would be value destroying strategies and are therefore not reasonable for a shareholder value incentivized management.

Although controlling for a battery of determinants, we cannot fully exclude the existence of unobserved variables. First, the areas *emission* and *resource reduction* as well as *workforce* and *society* are mutually highly correlated, implying that either of the areas could cause the effect or a combination of both. It is also conceivable that high scores in the *workforce*, *society*, *resource* or *emission reduction* areas serve as a proxy for some unobserved variable that may be the real cause, although the economic explanations support the causal influence of the mentioned areas. Firm fixed-effects models, that would overcome endogeneity problems caused by time-invariant variables, cannot be applied as they require the effect to be constant over time, while a change in the ESG scores or the area scores is related with a change in the management practice, which also could have an impact on the unobserved variables. Additionally, from a practical point of view, firm fixed effects would require more variability in the longitudinal section, i.e. a longer sample, which is currently unavailable since Asset4 commenced their rating business only in 2002. Nevertheless, independent of the exact causal relationship, from an asset pricing perspective our findings are promising. We provide evidence that high area scores predict future unanticipated cash flows that also lead to abnormal

returns. Furthermore, we demonstrate that the observed abnormal returns are not only due to an additional demand channel, which may be an investment fashion phenomenon that could change easily in the future.

Notes

1. For instance, reputation enhancement through good corporate citizenship is priced by stock markets efficiently (Brammer, Brooks, and Pavelin 2009).
2. We control the book-to-market portfolios for industry effects, as proposed by Wermers (2004).
3. This contrasts with many papers in financial economics that rely on the benchmark portfolio data for US stocks from Russ Wermer's homepage which rebalance the benchmark portfolios annually.
4. As a consequence, *return on assets*, which is another classical predictor, is omitted due to a high correlation with profitability.
5. As we have a relatively short panel and a broad cross section, this approach appears to be more suitable than classical panel regressions (Cameron and Trivedi 2010).
6. The standard deviations of the scores used for this sample are 30.91 (E), 28.54 (S), and 16.67 (G), which only deviate marginally from the figures reported in Table 1.
7. Note that we do not instrumentalize the ESG scores. From an investors perspective, we wish to investigate whether ESG scores can be used to forecast a persistent long-term abnormal return pattern, independent of the causal relation between ESG scores and abnormal returns.
8. To further investigate the long-term impact of the VS score, we compute cross-correlations of lagged VS scores with differentiated scores of all other areas. Indeed, lagged VS scores are positively correlated with all area scores in the environment and social dimension, respectively. Therefore, a high VS score is positively related to high E and S area scores in the medium term. These area scores may drive the observation.
9. We downloaded the respective risk factors from Kenneth French's website.
10. As a robustness check, we also computed the calender-time regression alphas with pure returns as endogenous variables. While the statistical interference is lower, the results are consistent with the results from Panel (a).

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Gregor Dorfleitner http://orcid.org/0000-0003-0270-0660
Sebastian Utz http://orcid.org/0000-0003-1570-752X
Maximilian Wimmer http://orcid.org/0000-0002-6283-3169

References

Aupperle, K. E., A. B. Carroll, and J. D. Hatfield. 1985. "An Empirical Examination of the Relationship between Corporate Social Responsibility and Profitability." *Academy of Management Journal* 28 (2): 446–463.

Barnea, A., and A. Rubin. 2010. "Corporate Social Responsibility as a Conflict between Shareholders." *Journal of Business Ethics* 97 (1): 71–86.

Bauer, R., N. Guenster, and R. Otten. 2004. "Empirical Evidence on Corporate Governance in Europe: The Effect on Stock Returns, Firm Value and Performance." *Journal of Asset Management* 5 (2): 91–104.

Bhagat, S., and B. Bolton. 2008. "Corporate Governance and Firm Performance." *Journal of Corporate Finance* 14 (3): 257–273.
Bhandari, L. C. 1988. "Debt/Equity Ratios and Expected Common Stock Returns: Empirical Evidence." *Journal of Finance* 43, 507–528.
Brammer, S., C. Brooks, and S. Pavelin. 2009. "The Stock Performance of America's 100 Best Corporate Citizens." *Quarterly Review of Economics and Finance* 49, 1065–1080.
Brennan, M. J., and P. J. Hughes. 1991. "Stock Prices and the Supply of Information." *Journal of Finance* 46, 1665–1691.
Brown, S. L. 1978. "Earnings Changes, Stock Prices, and Market Efficiency." *Journal of Finance* 33, 17–28.
Brown, L. D. 1997. "Analyst Forecasting Errors: Additional Evidence." *Financial Analysts Journal* 53, 81–88.
Cameron, A., and P. Trivedi. 2010. *Microeconometrics using Stata*. Revised ed. College Station, TX: Stata Press.
Chai, M. L., and S. Tung. 2002. "The Effect of Earnings-announcement Timing on Earnings Management." *Journal of Business Finance & Accounting* 29 (9-10): 1337–1354.
Chatterji, A., R. Durand, D. I. Levine, and S. Touboul. 2016. "Do Ratings of Firms Converge? Implications for Managers, Investors and Strategy Researchers." *Strategic Management Journal* 37, 1597–1614.
Chatterji, A., and D. Levine. 2006. "Breaking Down the Wall of Codes: Evaluating Non-financial Performance Measurement." *California Management Review* 48 (2): 29–51.
Chen, R. C. Y., S. W. Hung, and C. H. Lee. 2017. "Does Corporate Value Affect the Relationship between Corporate Social Responsibility and Stock Returns?" *Journal of Sustainable Finance and Investment* 7, 188–196.
Cheng, I.-H., H. G. Hong, and K. Shue. 2014. "Do Managers Do Good with Other Peoples' Money?" Fama-Miller Working Paper. Chicago Booth Research Paper No. 12–47.
Cheng, B., I. Ioannou, and G. Serafeim. 2011. "Corporate Social Responsibility and Access to Finance." Harvard Business School Working Paper 11(130).
Core, J. E., W. R. Guay, and T. O. Rusticus. 2006. "Does Weak Governance Cause Weak Stock Returns? An Examination of Firm Operating Performance and Investors' Expectations." *Journal of Finance* 61 (2): 655–687.
Cornell, B., and A. C. Shapiro. 1987. "Corporate Stakeholders and Corporate Finance." *Financial Management* 16 (1): 5–14.
Daniel, K., M. Grinblatt, S. Titman, and R. Wermers. 1997. "Measuring Mutual Fund Performance with Characteristic-based Benchmarks." *Journal of Finance* 52 (3): 1035–1058.
Daniel, K., and S. Titman. 1997. "Evidence on the Characteristics of Cross Sectional Variation in Stock Returns." *Journal of Finance* 52 (1): 1–33.
Deng, X., J. Kang, and B. S. Low. 2013. "Corporate Social Responsibility and Stakeholder Value Maximization: Evidence from Mergers." *Journal of Financial Economics* 110, 87–109.
Derwall, J., N. Guenster, R. Bauer, and K. Koedijk. 2005. "The Eco-efficiency Premium Puzzle." *Financial Analysts Journal* 61 (2): 51–63.
Derwall, J., and P. Verwijmeren. 2007. "Corporate Governance and the Cost of Equity Capital: Evidence from GMI's Governance Rating." ECCE Research Working Paper 06-01.
Desai, H., and P. C. Jain. 1997. "Long-run Common Stock Returns Following Stock Splits and Reverse Splits." *Journal of Business* 70 (3): 409–433.
Dhaliwal, D. S., O. Z. Li, A. Tsang, and Y. G. Yang. 2011. "Voluntary Nonfinancial Disclosure and the Cost of Equity Capital: The Initiation of Corporate Social Responsibility Reporting." *Accounting Review* 86 (1): 59–100.
Donelson, D. C., and R. J. Resutek. 2015. "The Predictive Qualities of Earnings Volatility and Earnings Uncertainty." *Review of Accounting Studies* 20 (1): 470–500.
Edmans, A. 2011. "Does the Stock Market Fully Value Intangibles? Employee Satisfaction and Equity Prices." *Journal of Financial Economics* 101 (3): 621–640.
Edmans, A., L. Li, and C. Zhang. 2014. "Employee Satisfaction, Labor Market Flexibility, and Stock Returns Around the World." NBER Working Paper 20300.

Fama, E. F., and K. R. French. 1988. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics* 22 (1): 3–25.
Fama, E. F., and K. R. French. 1992. "The Cross-section of Expected Stock Returns." *Journal of Finance* 47, 427–465.
Fama, E. F., and K. R. French. 2015. "A Five-factor Asset Pricing Model." *Journal of Financial Economics* 116 (1): 1–22.
Flammer, C. 2015. "Does Corporate Social Responsibility Lead to Superior Financial Performance? A Regression Discontinuity Approach." *Management Science* 61, 2549–2568.
Friede, G., T. Busch, and A. Bassen. 2015. "ESG and Financial Performance: Aggregated Evidence from More than 2000 Empirical Studies." *Journal of Sustainable Finance & Investment* 5 (4): 210–233.
Galema, R., A. Plantinga, and B. Scholtens. 2008. "The Stocks at Stake: Return and Risk in Socially Responsible Investment." *Journal of Banking & Finance* 32, 2646–2654.
Ghoul, S. E., O. Guedhami, C. C. Y. Kwok, and D. R. Mishra. 2011. "Does Corporate Social Responsibility Affect the Cost of Capital?" *Journal of Banking & Finance* 35 (9): 2388–2406.
Giroud, X., and H. M. Mueller. 2011. "Corporate Governance, Product Market Competition, and Equity Prices." *Journal of Finance* 66, 563–600.
Gompers, P., J. Ishill, and A. Metrick. 2003. "Corporate Governance and Equity Prices." *Quarterly Journal of Economics* 118 (1): 107–156.
Goss, A., and G. S. Roberts. 2011. "The Impact of Corporate Social Responsibility on the Cost of Bank Loans." *Journal of Banking & Finance* 35 (7): 1794–1810.
Gu, Z., and J. S. Wu. 2003. "Earnings Skewness and Analyst Forecast Bias." *Journal of Accounting and Economics* 35, 5–29.
Halbritter, G., and G. Dorfleitner. 2015. "The Wages of Social Responsibility – Where are They? A Critical Review of ESG Investing." *Review of Financial Economics* 26 (Supplement C): 25–35.
Hoepner, A., I. Oikonomou, B. Scholtens, and M. Schröder. 2016. "The Effects of Corporate and Country Sustainability Characteristics on the Cost of Debt: An International Investigation." *Journal of Business Finance & Accounting* 43 (1-2): 158–190.
Hong, H. G., J. D. Kubik, and J. A. Scheinkman. 2012. "Financial Constraints on Corporate Goodness." SSRN Working paper.
Humphrey, J. E., D. D. Lee, and Y. Shen. 2012. "Does it Cost to be Sustainable?" *Journal of Corporate Finance* 18 (3): 626–639.
Ioannou, I., and G. Serafeim. 2012. "What Drives Corporate Social Performance? The Role of Nation-level Institutions." *Journal of International Business Studies* 43, 834–864.
Kane, A., Y. K. Lee, and A. Marcus. 1984. "Earnings and Dividend Announcements: Is There a Corroboration Effect?" *Journal of Finance* 39 (4): 1091–1099.
Khan, M., G. Serafeim, and A. Yoon. 2016. "Corporate Sustainability: First Evidence on Materiality." *Accounting Review* 91 (6): 1697–1724.
Klock, M. S., S. A. Mansi, and W. F. Maxwell. 2005. "Does Corporate Governance Matter to Bondholders?" *Journal of Financial and Quantitative Analysis* 40 (4): 693–719.
Kothari, S. P., J. Lewellen, and J. B. Warner. 2006. "Stock Returns, Aggregate Earnings Surprises, and Behavioral Finance." *Journal of Financial Economics* 79, 537–568.
Luo, X., and C. B. Bhattacharya. 2006. "Corporate Social Responsibility, Customer Satisfaction, and Market Value." *Journal of Marketing* 70 (4): 1–18.
Luo, X., H. Wang, S. Raithel, and Q. Zheng. 2015. "Corporate Social Performance, Analyst Stock Recommendations, and Firm Future Returns." *Strategic Management Journal* 36 (1): 123–136.
Lyon, J. D., B. M. Barber, and C.-L. Tsai. 1999. "Improved Methods for Tests of Long-run Abnormal Stock Returns." *Journal of Finance* 54 (1): 165–201.
Margolis, J. D., H. A. Elfenbein, and J. P. Walsh. 2009. "Does it Pay to be Good … and Does it Matter? A Meta-analysis of the Relationship between Corporate Social and Financial Performance." SSRN Working Paper.
Orlitzky, M. 2008. "Corporate Social Performance and Financial Performance: A Research Synthesis." In *The Oxford Handbook of Corporate Social Responsibility*, edited by A. Crane, A. McWilliams, D. Matten, J. Moon, and D. S. Siegel, 113–134. Oxford University Press.

Orlitzky, M., F. L. Schmidt, and S. L. Rynes. 2003. "Corporate Social and Financial Performance: A Meta-analysis." *Organization Studies* 24 (3): 403–441.
Peyer, U., and T. Vermaelen. 2009. "The Nature and Persistence of Buyback Anomalies." *Review of Financial Studies* 22 (4): 1693–1745.
Pfarrer, M. D., T. G. Pollock, and V. P. Rindova. 2010. "A Tale of Two Assets: The Effects of Firm Reputation and Celebrity on Earnings Surprises and Investors' Reactions." *Academy of Management Journal* 53, 1131–1152.
Pontiff, J., and L. D. Schall. 1998. "Book-to-Market Ratios as Predictors of Market Returns." *Journal of Financial Economics* 49, 141–160.
Posner, B. Z., and W. H. Schmidt. 1992. "Values and the American Manager: An Update Updated." *California Management Review* 34 (3): 80–94.
Rendleman, R. J., C. P. Jones, and H. A. Latane. 1982. "Empirical Anomalies based on Unexpected Earnings and the Importance of Risk Adjustments." *Journal of Financial Economics* 10, 269–287.
Schwert, G. W. 1983. "Size and Stock Returns, and Other Empirical Regularities." *Journal of Financial Economics* 12, 3–12.
Stapleton, R. C., and M. G. Subrahmanyam. 1977. "Market Imperfections, Capital Market Equilibrium and Corporation Finance." *Journal of Finance* 32 (2): 307–319.
Statman, M., and D. Glushkov. 2009. "The Wages of Social Responsibility." *Financial Analysts Journal* 65 (4): 33–46.
Stellner, C., C. Klein, and B. Zwergel. 2015. "Corporate Social Responsibility and Eurozone Corporate Bonds: The Moderating Role of Country Sustainability." *Journal of Banking & Finance* 59, 538–549.
US SIF. 2014. 2014 report on sustainable and responsible investing trends.
von Wallis, M., and C. Klein. 2015. "Ethical Requirement and Financial Interest: A Literature Review on Socially Responsible Investing." *Business Research* 8 (1): 61–98.
Watts, R. L. 1978. "Systematic 'Abnomal' Returns after Quarterly Earnings Announcements." *Journal of Financial Economics* 6, 127–150.
Wermers, R. 2004. "Is Money Really 'Smart'? New Evidence on the Relation between Mutual Fund Flows, Manager Behavior, and Performance Persistence." SSRN: https://protect-us.mimecast.com/s/dqYaBRiL6YgmFr?domain=ssrn.com.
Wurgler, J., and E. Zhuravskaya. 2002. "Does Arbitrage Flatten Demand Curves for Stocks?" *Journal of Business* 75 (4): 583–608.