Corporate Social Responsibility: How much is enough? A Higher Dimension Perspective of the Relationship between Financial and Social Performance

Kalaitzoglou, I., Pan, H. & Niklewski, J.

Original citation & hyperlink:
Kalaitzoglou, I, Pan, H & Niklewski, J 2021, 'Corporate Social Responsibility: How much is enough? A Higher Dimension Perspective of the Relationship between Financial and Social Performance', Annals of Operations Research, vol. 306, pp. 209-245.
https://doi.org/10.1007/s10479-020-03834-y

DOI 10.1007/s10479-020-03834-y
ISSN 0254-5330
ESSN 1572-9338

Publisher: Springer

The final publication is available at Springer via http://dx.doi.org/10.1007/s10479-020-03834-y

Corporate Social Responsibility: How much is enough?
A Higher Dimension Perspective of the Relationship between Financial and Social Performance

Iordanis Kalaitzoglou[1]
Audencia Nantes
Business School

Hui Pan[2]
Coventry Business School
Coventry University

Jacek Niklewski[3]
Coventry Business School
Coventry University

Abstract

We investigate the nature of the relationship between Corporate Social Responsibility (*CSR*) and Corporate Financial Performance (*CFP*) by examining how it changes across a third dimension that accounts for firm-specific factors. We propose a semi-latent specification of an endogenous control variable, which can, for the first time, explicitly identify, for each individual firm, the threshold level where the marginal impact of *CSR* on *CFP* turns positive. We provide empirical evidence that this threshold depends on the additional dimension and consequently, the previously reported U-shape seems to be an aggregation of relationships of differential magnitude and direction. This disaggregation fits the data better and therefore, we maintain that the addition of a higher dimension, along with the identification of the threshold level, can explain the conflicting results in the literature.

Keywords: *Corporate Social Responsibility (CSR), Firm Size, Corporate Financial Performance (CFP), Endogeneity, Asymmetric Relationship.*

---

[1] Corresponding author. Address: 8 route de la Jonelière - B.P. 31222 44312, France. e-mail: ikalaitzoglou@audencia.com, tel: +33 (240) 378402.

[2] Priory Street, Coventry CV1 5FB, United Kingdom. E-mail: H.Pan@coventry.ac.uk, tel: +44 24 7765 7755.

[3] Priory Street, Coventry CV1 5FB, United Kingdom. E-mail: J.Niklewski@coventry.ac.uk, tel: +44 24 7765 8953.

1 Introduction

"Does it pay to be good?" (Trudel and Cotte 2009). This question refers to the marginal impact of Corporate Social Responsibility (*CSR*) on Corporate Financial Performance (*CFP*). The initial response of the literature provides a full spectrum of arguments that range from a positive (Jones 1995) to negative (Aupperle et al. 1985) relationship. These inconclusive propositions have led to the realization that *CSR* might not always be profitable and therefore, a more appropriate question is not "whether", but "under what conditions" does it pay to be good (Rowley and Berman 2000). Driven by the absence of a strictly monotonic *CSR*-*CFP* relationship, several studies suggest that these conflicting results can be attributed to an endogenous (Waddock and Graves 1997) or an asymmetric (e.g., Barnett 2007) relationship. This modelling of higher degrees of variation suggests that a binary approach might be limiting in capturing the complexity of the *CSR*-*CFP* nexus (Crane et al. 2018) and implicitly assumes that there might be another level of interaction, governed by observable or unobservable factors.

Addressing this concern, a branch of literature suggests the existence of observable 'mediators' (e.g., Saeidi et al. 2015) or 'moderators' (e.g., Gully et al. 2013), or in a wider scope multiple criteria (e.g., Lamata et al. 2018) or different 'management recipes' (e.g., Isaksson and Woodside 2016), which might affect the way *CSR* interacts with *CFP*. They recognize explicitly that the exact relationship might also depend on other factors, but their approach focuses on identifying whether they might affect (magnitude) the *CSR*-*CFP* relationship rather than what this relationship might look like (shape). Other studies try to describe the shape of the *CSR*-*CFP* link by considering the existence of non-linearities that might arise due to the interaction with latent (unobservable) factors, such as the ability of the firm to interact with its stakeholders (Barnett 2007) or the ambiguity about the quality of a project (Flammer 2015). Because these latent factors cannot be identified/measured, the modelling needs to explicitly describe the shape of the *CSR-CFP* link by imposing a structural form, usually a quadratic function, on *CFP*. Consequently, the literature operates under an implicit trade-off between the functional form (shape) and the interacting factor (latent/observable).

We directly address this trade-off by introducing an additional dimension, which accounts for firm specific factors, to the *CSR-CFP* relationship, and we explicitly model how this dimension might affect its shape. This way, we can, for the first time, identify for each individual firm the exact minimum level (threshold) of *CSR* required for its impact on *CFP* to turn positive and thus, the shape of the curvature at firm-specific (rather than at sample) level.

We conduct a global empirical analysis which suggests that the previously reported U-shape cannot be unconditionally applied to all firms. In contrast, we report that the threshold value depends on a third dimension and therefore, we describe, for the first time, the exact intensity of curvature at a firm-specific, rather than at a sample, level. Consequently, we maintain that the overall U-shaped *CSR-CFP* link is observed across sample due to an aggregation of differently shaped links across a higher dimension. The identification for the inflection point is robust to alternative control variables, as well as to measurement/functional form/cross-sectional differences. Consequently, our approach can explain previously conflicting results that arise due to sampling, because it focuses on a more fundamental (firm-specific) level.

2 Nature of the *CSR-CFP* Relationship

"Doing good and doing well" (Avi-Yonah 2005) or "Does it pay to be good?" (Trudel and Cotte 2009) are two versions of the same question that literature has been concerned with over the last half a century, referring to whether the marginal impact of *CSR* on *CFP* can be positive.

2.1 Endogeneity

Early approaches, driven by the shareholder theory (e.g., Friedman 1970) see *CSR* as a "donation" and thus as a wealth transfer from shareholders to a larger group of stakeholders.

This leads to diminishing wealth and to increasing agency problems (e.g., Jensen and Meckling 1976). Consequently, early literature suggests a negative impact of *CSR* on *CFP* (Aupperle et al. 1985). In contrast, a stakeholder theory-related approach (Freeman 1984) suggests that social externalities could be traced back to the firm; therefore, they could affect shareholders' wealth (Agle et al. 2008). Consequently, the goal of a firm that actively tries to better meet its stakeholders' needs is to increase its shareholders' wealth (e.g., Jones 1995); thus, the impact of *CSR* on *CFP* should be positive (e.g., Li et al. 2018). Empirical and theoretical literature confirms the positive link and attributes this to factors such as enhanced morale, efficiency, and/or productivity (Solomon and Hanson 1985; Barney 1991; Russo and Fouts 1997; Porter and Kramer 2006).

Trying to reconcile these two opposing views, several studies focus on the direction of the relationship and potential endogeneity. Relevant literature reports that companies with higher *CFP* tend to invest more in *CSR* (Hillman and Keim 2001; Orlitzky et al. 2003); therefore, no safe conclusion can be reached unless the reverse causality is also addressed. Endogeneity is very relevant in our study because its existence would explicitly challenge the "universal" character of the *CSR*-*CFP* relationship. This is also consistent with the propositions of Bénabou and Tirole (2010) who argue that the existence of endogeneity makes inference almost impossible. They even suggest that if endogeneity is quantitatively treated, primarily with instrumental variables (e.g., Shahzad and Sharfman 2017), the intensity of the link is significantly diminished (e.g., Waddock and Graves 1997; Garcia-Castro et al. 2010). We explicitly recognize the importance of endogeneity and suggest a structural equation approach that lets the data identify the existence and magnitude of endogeneity. This approach does not simply address endogeneity by simply "treating" the data, it explicitly models it allowing the data to identify and measure it.

2.2 Asymmetry

Other studies attribute the differences to asymmetries. Brammer and Millington (2008) propose that the positive association between *CSR* and *CFP* follows diminishing and decreasing returns. Consequently, if the scope of social responsibility participation strays beyond management in addressing social concerns (e.g., with little or no effect on stakeholders), the net effect is likely to be declining financial performance. They also argue that the correlation between *CSR* and performance is highest at the extremes, showing that financial performance is high at both very high and very low *CSR* levels. Barnett (2007) argues that the U-shaped relationship depends on the firm's ability to "better meet" stakeholders' needs and thus capitalize on *CSR* investments, a concept referred to as Stakeholder Influence Capacity (SIC). Barnett and Salomon (2012) report that firms with low/high *CSR* exhibit higher *CFP* than firms with moderate *CSR*. Along the same lines, Flammer (2015) and Flammer and Bansal (2017) identify the existence of threshold effects due to ambiguity. Using a discontinuity approach, marginal *CSR* investments in the sense that there is increased ambiguity about their future outcome as it is reflected on a "close-call" acceptance/rejection, are found to be significantly different from their decreased ambiguity counterparts. These studies explicitly suggests that not all *CSR* projects should be expected to have a monotonic impact on *CFP*, indicating an optimal *CSR* intensity (e.g., Jawahar et al. 2015) or else a threshold/inflection point (e.g., Barnett and Salomon 2012) where the marginal impact of *CSR* on *CFP* changes. Our approach enables a data-driven identification of the shape and magnitude of these asymmetries, as well as of the exact inflection point at a firm-specific level.

2.3 Additional Dimension

This asymmetry, attributed mostly to latent (unobservable) underlying factors (e.g., SIC or ambiguity), involves various aspects that are of relevance to our study. First, it recognizes explicitly that the intensity and shape of the *CSR*-*CFP* link might depend on other factors. This resonates with the view of other studies (Galbreath and Shum 2012) suggesting that the *CSR*-

CFP relationship is determined by mediators or moderators, such as customer satisfaction, reputation, and competitive advantage (e.g., Walsh and Beatty 2007), or even the capital structure (e.g., Cornett et al. 2016), or a combination of them (e.g., Lamata et al. 2018) into a 'managerial recipe' (e.g., Isaksson and Woodside 2016). The 'mediators/moderators/multi-criteria' approach investigates the *CSR-CFP* link beyond a binary context but focuses only on identifying whether one or more observable factors affect the *CSR-CFP* link without addressing its shape or its magnitude. In contrast, other studies employ unobservable factors, such as SIC (Barnett 2007) or "ambiguity" (Flammer 2015) in order to encompass a multitude of interacting factors, but because of the inability to measure them, they need to imply a functional form – mostly quadratic (Barnett and Salomon 2012) or with threshold effects (Flammer and Bansal 2017). This approach might be prone to misspecification error, while it is also invariant across the underlying factor.

Second, if a factor makes the *CSR-CFP* link to follow a pattern, then the intensity of a *CSR* strategy should also depend on this factor. For example, Barnett (2007) suggests that a more intense *CSR* strategy should be preferred when SIC is high, and vice versa. However, the literature implies a trade-off between identifying either the interacting factor or how it affects the *CSR-CFP* link (functional form), therefore, it cannot be empirically tested on a firm level.

We recognize that the *CSR-CFP* link should be examined beyond the binary nexus and that an explicit modelling of a third dimension is necessary to capture the complexity of the relationship. We propose a structural modelling of the additional dimension as an observable endogenous variable, with which we can identify uniquely per firm the threshold level at which the impact of *CSR* performance on *CFP* turns positive.

2.4 Reflections on Previous Literature

Collectively, previous literature on the *CSR-CFP* link has evolved from approaching the impact of *CSR* on *CFP* as a one-dimensional binary problem to accepting that it is a multidimensional issue with distinct sampling properties. There seems to be a consensus on the fact not all *CSR* activities are profitable and, therefore, different samples exhibit different shapes and magnitude of the *CSR-CFP* link. Previous studies that try to explain these differences attribute them to potential endogenous and asymmetric effects. More recent literature, which follows a more intensive quantitative approach (e.g, Crane et al. 2018), attempts to model these asymmetries and account for the endogeneity, but we understand that it operates under a significant trade-off; either explicitly defining how the relationship between *CSR* on *CFP* looks like, i.e., describing precisely its structural form, or considering that it is governed by latent factors. Consequently, previous models define either the factors or the structural form, but not both.

We purport that this trade-off might be the reason why several empirical studies exhibit cross-sectional differences and report conflicting results. We extend previous literature by addressing this trade-off directly. We do so by proposing a new modelling that investigates potential endogeneity and asymmetries in data driven way, without a pre-specification of the structural form. This is done by introducing and modelling an (extendable) additional dimension, the magnitude of which affects the shape and the magnitude of the *CSR-CFP* link. Different levels of the additional dimension exhibit a different degree of endogeneity and asymmetry, varying from high convexity to high concavity. Each firm operates at a particular level of the additional dimension, which is linked to a particular shape of the *CSR-CFP* link (i.e., degree of curvature/concavity). Consequently, our model can estimate the exact degree of curvature of the *CSR-CFP* relationship at a firm, rather than at an aggregate level. Previous studies present cross-sectional results aiming at describing the overall shape of the *CSR-CFP* relationship. We decompose this cross-sectional estimate across a third dimension, which explicitly models how this overall/cross-sectional relationship, reported in previous studies, is composed from each individual firm.

We consider this shift from an aggregate level to a firm specific estimate as our major contribution that has the potential to capture various stylized factors (e.g., Bass and Milosevic 2018) in a more unified context and thus, explain the conflicting findings in the literature. The empirical literature reports regional (Shahzad and Sharfman 2017) or cross-sectional (Rowley and Berman 2000) inconsistencies, which, we suggest, emerge from inability to identify the shape of the *CSR-CFP* link on a firm level. Our approach shifts the focus from a general/universal modelling to a more granular one, where cross-sectional differences – captured by the intensity of the additional dimension – are treated as a determinant of its shape, rather than as inconsistencies.

3 Methodology

3.1 Data

In our analysis we employ three different datasets to investigate the relationship between *CSR* and *CFP*, especially focusing on a potentially higher order interaction. First, we employ a primary sample with global coverage, based on "Vigeo" *CSR* rating, to test whether an additional, observable, dimension can adequately capture different cross-sectional/regional differences. [4] In addition, we employ two secondary samples for testing the robustness of our findings to different sampling properties and *CSR* valuation methods. They are based on the constituents of the S&P1500 and S&P500 index with valid data for two periods; 1997-2010 (pre-2010) and 2011-2017 (post-2010), using the "KLD" (pre-2010) CSR rating and the Environmental, Social and Governance (ESG; post-2010) performance index provided by Bloomberg, respectively.[5] For more information please refer to the online appendix.

[4] *CSR* is the discretionary societal expectations of a firm (Carroll 1979). This explicitly differentiates social from financial objectives, implicitly recognizing that not all *CSR* actions are profitable. However, because this is a forward-looking (latent) concept, we use an empirical proxy, Corporate Social Performance (CSP) (Carroll et al. 2016), which measures how socially responsible the firm has been in the past. This backward-looking proxy does not introduce estimation bias because we employ a backward-looking measure of *CFP*.

[5] KLD methodology changed due to the transition to the MSCI ESG indices family that occurred on the 1st of September 2010; therefore, we collect KLD data until 2010 and the ESG (Bloomberg) data beyond 2010 to

The primary dataset employed consists of all firms rated by "Vigeo" between 1997 and 2012, according to their *CSR* performance. Vigeo provides *CSR* performance scores $0 \leq H_m \leq 100$ on six domains, $m = 1, \dots, M = 6$, which are used to construct a weighted score $CSR = \frac{1}{100xM}\sum_{m=1}^{M} H_m$, equivalent to their net social performance (Chatterji et al. 2009).[6]

We cleaned the data by dropping all observations reporting negative equity capital. To match financial data with *CSR* information, only the most up-to-date annual ratings were considered. Furthermore, all observations outside a 5σ-confidence interval were considered outliers and thus omitted. This results in a pooled dataset with a total of 7,032 firm-year observations. We split the sample into three broad regional sub-samples, namely the United States (US), Europe (EU), and Rest Of the World (ROW), implicitly assuming different market-stylized factors, such as market maturity and risk. To account for industry-specific effects, we use the Thomson Reuters (EIKON) classification, which is also the source for financial firm-specific data.

In order to address previous concerns regarding *CSR* measurement (Carroll et al. 2016), we test the robustness of our methodology on the KLD data and the ESG disclosure index provided by Bloomberg.[7] In addition to the Vigeo dataset, we employ a second sample that consists of all firms included in the S&P1500 index with valid observations between 1997 and 2010 and add the corresponding KLD data and industry codes to the sample.[8] We aggregate the strengths and weaknesses of each firm to create its net social performance score (Barnett and Salomon

---

maintain consistency. Further information about the exact methodologies and social performance valuation can be found in RiskMetrics Group (2010) for KLD and in Bloomberg (2020) for ESG.

[6] The data collection focuses on the period before the formalization of the Vigeo Euronext Indices and the Merge with EIRIS. Further information can be found in Vigeo Eiris (2020).

[7] Carroll et al. (2016) argue that CSP scores might not reflect the true CSP of the firms, due to biases, such as regional factors (Shahzad and Sharfman 2017), the weighting used (e.g., García-Melón et al. 2016; Capelle-Blancard and Petit 2017) or simply because they do not adequately cover the breadth of *CSR* (e.g., Lamata et al. 2018; Oll et al. 2018). Even worse, Chatterji et al. (2016) report that the existing CSP scores do not converge and therefore, they might not be a good proxy for *CSR*. We address these concerns by testing the robustness of our findings to CSP-scores with different criteria, weighting and regional/cross-sectional coverage (Vigeo, KLD and ESG), which should be able to capture the diversity of these concerns.

[8] The coverage of the KLD index was not extensive in the early stages and therefore, this sample mainly consists of the constituents of the S&P500 index. The estimation results are consistent when focusing only on S&P500.

2012). KLD rates the firms based on 13 individual social performance criteria and the scores given are +1 for a strength and -1 for a weakness/concerns. The score is computed as $CSR_{it} = \sum_{m=1}^{n} KLD_{mit}$, where KLD_m is the score (-1 or +1) of category m and its range is from -12 to 15. We apply the same data thinning process, which results in a pooled dataset of 10,866 firm-year observations.

Finally, in order to test the robustness of our findings to a different *CSR* score, as well as to intertemporal differences, we employ a third (final) dataset, which consist of all firms of the S&P500 index with valid observations between 2011 and 2017 and add the corresponding ESG data and industry codes to the sample. Bloomberg reports an aggregated ESG score with a range from 0 to 100 and we compute our *CSR* score as $CSR_{it} = ESG_{it}/100$. We apply the same data thinning process, which results in a pooled dataset of 3,060 firm-year observations.

3.2 Model

The proposed model can be summarized in the following system of simultaneous equations:

$$\begin{cases} CFP_{it} = \left(\alpha_0 + \sum a_{0,p} I_{pit}\right) + \alpha_1 Z_{it} + a_2^{'} CSR_{it} + \sum \alpha_q CV_{CFP,qit} + \varepsilon_{1it}, & (1) \\ CSR_{it} = \left(\beta_0 + \sum \beta_{0,p} I_{pit}\right) + \beta_1 CFP_{it} + \beta_2 Z_{it} + \sum \beta_q CV_{CSR,qit} + \varepsilon_{2it}, & (2) \\ Z_{it} = \left(\mu_0 + \sum \mu_{0,p} I_{pit}\right) + \mu_1 CFP_{it} + \mu_2 CSR_{it} + \sum \mu_q CV_{Z,qit} + \varepsilon_{3it}, & (3) \end{cases}$$

where $A = (a_1, \dots, a_q)$, $B = (\beta_1, \dots, \beta_q)$, $\Gamma = (\gamma_0, \gamma_1)$, and $U = (\mu_1, \dots, \mu_q)$ are vectors of parameters to be estimated, and $p = (Industry, Country, Year)$ is a vector of dummy variables identifying a firm $i = 1, 2, \dots, n$ at time t.[9] CFP_{it} is the corporate financial performance, captured by return on equity, $ROE_{it} = Net\ Income_{i,t}/Equity_{i,t}$. Z_{it} is how we introduce an additional dimension in the form of a potentially endogenous variable. Z is an

[9] The model can be estimated with various estimation methods appropriate for a system of equations. We use an iterative GMM procedure with Newey–West heteroskedasticity–consistent errors, as our main estimation method, but we also test the robustness of our findings with simpler estimation methods and we find them to be consistent. The use of instrumental variables can account for various econometric issues, identify more precisely the CSR-CFP link, while it is also consistent with the literature. For more information, please refer to the online appendix.

observable variable, whose specification is modelled in Eq. (3). However, it appears in Eq. (1) and (2) as a semi-latent specification (Lewbel 1998) and this way it can indirectly capture other firm specific factors. This is how our study tries to bridge the trade-off between functional form (Eq. (1)) and identification of the interacting variable (Eq. (3)).

Furthermore, Eq. (1) investigates a potential asymmetric impact of *CSR* on *CFP*. Coefficient $a_2^{'}$ captures the impact of *CSR* on *CFP* and is dissected into two regimes: high and low, per Eq. (4) which is a logistic smooth transition function (Chan and Tong 1986; Van Dijk et al. 2002) with a fixed smoothness parameter (i.e., equal to 1):

$$a_2' = a_2^{low}\left(\frac{1}{1+e^{\left(\frac{CSR_{it}-\underbrace{\{\gamma_0+\gamma_1 Z_{it}\}}_{s_{it}}}{}\right)}}\right) + a_2^{high}\left(1-\frac{1}{1+e^{\left(\frac{CSR_{it}-\underbrace{\{\gamma_0+\gamma_1 Z_{it}\}}_{s_{it}}}{}\right)}}\right), \tag{4}$$

where a_2^{low} (a_2^{high}) captures the marginal impact of *CSR* on *CFP*, when $CSR_{it} < s_{it}$ ($CSR_{it} > s_{it}$). If they are different and statistically significant, this would indicate that the impact of *CSR* on *CFP* is different when *CSR* is below or above the threshold (asymmetric). s_{it} is a threshold value, which determines the *CSR* level that changes its marginal impact on performance and it is identified uniquely for every i and *t*. Depending on the estimates of the parameters, the threshold value s_{it} might be constant across the sample (i.e., γ_0) or might depend on Z (i.e., γ_1). The sign, the magnitude and the significance of these coefficients provide a full description of the *CSR*-*CFP* link, as well as how and how much it is affected by the additional dimension. The exponential function allows for a smooth asymmetric effect, the degree of which depends on the magnitude of *Z*. This allows for infinite variations, one for every firm, of the shape of the relationship between *CSR* and *CFP*, and, as such, this specification is less susceptible to misspecification error. If a_2^{low}, a_2^{high}, and γ_1 are statistically significant, our model would generate a differently shaped relationship between *CSR* and *CFP* across different levels of Z. Eq. (4) is the novelty that our approach suggests and it describes how a dimension/variable Z can affect the inflection point on each firm, as well the degree of curvature, a_2^{low} and a_2^{high}.

Furthermore, the latent character of Z (Eq. (3)) implicitly assumes that other factors, ε_{3t}, not included in the vector $CV_{Z,qit}$, can affect the shape of the relationship between *CSR* and *CFP* in a non-linear fashion (Eq. (4)). This increases the generality of the threshold variable, Z, while reducing the omitted variable misspecification error. The specification in Eq. (4) can be easily expanded to accommodate more threshold variables, fixed effects and/or a higher degree of asymmetry and thus account for a wider variety of firm-specific factors.

3.3 Model Flexibility and Potential Extensions

One of the major contributions of our model over previous literature is that it is 'investigative' in nature, without imposing any predisposition with respect to the direction, shape or magnitude of the relationship between *CSR* and *CFP*. Previous models impose a structural form, motivated by an existing concept, e.g., the impact of a moderator/mediator, and test whether it is valid on a dataset or not. Our approach here is a lot more general and provides a way of accommodating various concepts. More precisely, we do not make any prior assumptions with respect to the direction/structure of the *CSR-CFP* relationship, but we let the data determine this relationship. We do that by using a system of equations, instead of a specific structural form and/or data manipulation that addresses endogeneity or sampling issues, which according to the sign and significance of the coefficients can identify the existence of i) endogeneity, ii) direction, iii) mediators, iv) moderators and/or v) a combination of them in a completely data driven way. Consequently, our model and its potential extensions that are discussed below could be used to test various theories or sampling properties without imposing a structural form or specific data treatment.

In more detail, different combinations of coefficient signs and significance would indicate the presence of various, previously reported, aspects of the *CSR-CFP* relationship. For example, in Eq. (1) and Eq. (2), coefficients $a_2^{'}$ and β_1 investigate whether there is any endogenous relationship between *CSR* and *CFP*. If only one of the two is significant, our model identifies that there is a one-directional effect, while if both are significant it identifies the presence of endogeneity. This is done in a natural way, without altering the structural form or the dataset and therefore if provides a general way of testing directional effects. This can also be extended to investigate endogeneity of higher order. Coefficient β_2 captures the impact of *Z* on *CSR*, while Eq. (3) examines potential endogeneity among *CSR*, *CFP*, and Z as captured by coefficients μ_1 and μ_2 . This approach is preferred over an instrumental variables specification (e.g., Cornett et al. 2016) because it introduces a level of generality that can conceptually nest/empirically test previous approaches. For example, one of the most important aspects of our model is that it can test for higher order interactions and their nature, e.g., moderators or mediators, in a natural data driven way. If only $a_2^{'}$ is found to be significant, while α_1 remains insignificant, this would indicate that the control variable used as a third dimension is a pure moderator, while if all α_1, $a_2^{'}$ and μ_2 are statistically significant, then *CSR* has both a direct and an indirect impact on *CFP*, with *Z* acting as a mediator. Furthermore, the composition of Eq. (3) could capture a 'managerial recipe' (Isaksson and Woodside 2016), especially if *Z* is defined as a completely latent variable.

Another major contribution of our model is the introduction of Eq. (4), which is flexible enough to capture various shapes of the link between *CSR* and *CFP*, without imposing a specific structural form like a quadratic function (e.g., Barnett and Salomon 2012) or structural breaks (e.g., Flammer and Bansal 2017). In more detail, Eq. (4) is flexible enough to allow a linear (e.g., monotonically increasing (decreasing) $a_2^{low} = a_2^{high} > 0$ ($a_2^{low} = a_2^{high} < 0$)) or a non-linear (e.g., exponential growth (decay) $0 > a_2^{low} > a_2^{high}$ ($a_2^{low} < a_2^{high} < 0$) and logistic growth (decay) $0 > a_2^{high} > a_2^{low}$ ($a_2^{high} < a_2^{low} < 0$)) relationship. It can also accommodate a concave (e.g., $a_2^{low} < 0$ and $a_2^{high} > 0$) shape and a convex (e.g., $a_2^{low} > 0$

and $a_2^{high} < 0$) shape. The logistic function simply assumes exponential smoothing and does not affect the fundamental shape of the relationship.

However, the selection of an observable variable as a threshold variable might be limiting, in the sense that it might not be able to capture non-observable factors. To this extent, the inclusion of ε_{3t} adds a latent character to the threshold variable Z. This implies that factors not considered in Eq. (3) can also influence the shape of the relationship between *CSR* and *CFP*; therefore, it implicitly assumes that it is not only Z that affects this relationship but other relevant factors (e.g., known ($CV_{Z,i,t}$) and unknown (ε_{3t}) factors) through their impact on Z. Consequently, this specification allows a threshold variable to be used as an empirical proxy for the status of the firm.

If this is not sufficient, though, or if it is too restrictive, the model can be easily extended in various directions. If more firm-specific factors need to be considered, Eq. (4) can be extended by allowing more variables to affect the threshold values (e.g., $s_{it} = \gamma_0 + \sum_{j=1}^{J} \gamma_j TV_{jit}$) where TV is a vector of threshold variables and/or fixed effects. Furthermore, if focus lies on the state of the relationship (e.g., high or low) rather than on the shape or the degree of the transition, a more stochastic approach could be selected. Instead of using a deterministic (e.g., size) or semi-latent (e.g., structural equation) variable, a completely unobservable (latent) variable with observable discrete states could be employed, where the transition from a_2^{low} to a_2^{high} follows a Markov switching framework. Finally, a greater number of regimes could be considered, should a higher degree of non-linearity be required.

3.4 Introducing a Higher Dimension

In this study, we claim that examining the *CSR-CFP* link in a binary context is inadequate to capture firm, market, or regional stylized factors; therefore, empirical studies so far generate conflicting results. Instead, we share the view that the *CSR-CFP* link might be affected by other mediators/moderators or latent factors. In the model presented above we propose and explicitly model an additional dimension of the *CSR-CFP* link, which requires the selection of an observable variable as an intermediating factor. For this purpose, we choose firm size ($Size_{it} = log(Total\ Assets_{i,t})$), as one of the possible options, because it is well reported in the literature to be highly correlated with both *CSR* and *CFP* (e.g., Bowen 2002). This is not an identification statement and we also test the robustness of our findings with alternative firm specific factors, such as $R\&D$ expenses.[10]

Unlike early literature (e.g., Mansfield 1962) and the *'Law of Proportionate Effect*' (Gibrat 1931), several studies report a significant interaction between firm size and *CFP*. A branch of literature (e.g., Shepherd 1972) suggests that larger firms enjoy increased financial returns due to economies of scale (e.g., Stigler 1958), higher efficiency (Rappaport 1998), greater market share (Amato and Wilder 1985), and/or greater market power/concentration (Shepherd 1972), although they suggest that the exact relationship exhibits strong industry effects (e.g., Amato and Amato 2004). In contrast, several other studies (e.g., Evans 1987) report a negative correlation between firm size and profitability due to diseconomies of scale (e.g., Ratchford and Stoops 1998), diminishing returns to the fixed productivity factors (e.g., Marshall 1961), and/or organizational costs (e.g., Williamson 1967). In parallel, other studies argue that the link between firm size and *CFP* is highly empirical (Audretsch et al. 2002) and might depend on the balancing of administrative overheads and fixed costs, which decrease with size but increase with organizational complexity (Blau 1970).

[10] The selection of firm size as a third dimension is by no means an identification statement. We recognize that a single variable might not be enough to capture the complexity of the *CSR-CFP*. Therefore, we suggest a semi-latent modeling to account for other factors too, as well as various model extensions in order to incorporate other factors of interest or other functional forms. Consequently, by selecting firm size, we do not claim that it is the only or the best factor affecting the *CSR-CFP* link. Instead, we select a variable that is recognized in the literature to affect both, aiming at highlighting how an additional dimension can help explaining their relationship.

Along the same lines, a significant part of the literature recognizes the impact of firm size on *CSR*. Early literature, driven by agency theory, links the intensity of *CSR* to greater managerial autonomy (Atkinson and Galaskiewicz 1988) and utility (Navarro 1988), which are more evident in larger firms. The stakeholder theory complements this view by focusing on the economic benefit of *CSR*. Firms that better meet their stakeholders' needs are rewarded by higher legitimacy (Hooghiemstra 2000) and better access to resources. Inevitably, the literature at this stage recognizes that organizational (i.e., firm-specific) characteristics might affect the outcome of *CSR* actions (e.g., Pfeffer and Salancik 1978). The firm's organizational architecture is industry dependent, but it is also highly correlated to firm size; therefore, it is seen as a major determinant of the effectiveness of *CSR* actions (Udayasankar 2008).

Firm size is linked to market power, visibility, and governance structures, which are hypothesized to have a strong impact on *CSR* strategy implementation (e.g.,Etzion 2007; Aguinis and Glavas 2012). Although some studies (e.g., Meznar and Nigh 1995) consider larger firms more resistant to external influences and thus less socially responsive, most studies consider larger firm size to be associated with more intense *CSR* strategies. They consider larger firms to have better established governance (Schreck and Raithel 2018) and administrative practices (Donaldson 2001), and this would lead to a greater responsiveness to social issues (Brammer and Millington 2005). In addition, larger firms are more visible (e.g., Etzion 2007; Brammer and Millington 2008), and it should be easier for them to convey information to their stakeholders (Darnall et al. 2010). They tend to invest more in *CSR*, especially in the presence of an intrinsic value that increases their competitive advantage (Chih et al. 2010), or due to higher corporate reputation (e.g., Fombrun and Shanley 1990), or in order to reduce the asymmetry of information between managers and shareholders (Brammer and Millington 2005). In parallel, their investment strategy should be affected to a greater degree by shifts in stakeholders' needs, as it is easier to associate them with "good" or "bad" practices, due to increased visibility (Watts and Zimmerman 1986), increased social pressure (Aguilera et al. 2007), or reduced size-related costs because of *CSR* disclosure (Ness and Mirza 1991).

Consequently, size is found to be a significant determinant of both *CSR* and *CFP*, with implications as for SMEs or non-listed firms (e.g., Spence 2016). We build on this idea and propose that, since size interacts with both, it might be endogenous (Orlitzky 2001; Surroca et al. 2010) with either/both and that it might affect the effectiveness of *CSR* actions and thus the way *CSR* and *CFP* interact. We try to merge all these concepts by considering firm size as an integral part of the *CSR*-*CFP* nexus. Again, we stress out that this is not an identification statement. Firm size is selected because it has been reported in the literature to have an impact on both *CSR* and *CFP* and potentially on both or on how they interact. Consequently, we consider it a suitable starting point to be investigated as a third dimension. Our framework could empirically identify how it interacts with *CSR* and *CFP*, in a data driven way, without imposing any functional form or conceptual restriction.

3.5 Confounding Effects and Control Variables

Our model provides a general setup that the employed in order to investigate various aspects of the *CSR*-*CFP* relationship. However, this cannot be done without accounting for confounding effects. Our system of equations approach is very flexible and allows for a selection of different confounding factors for each one of the main components of the model, namely *CSR*, *CFP* and *Z*. Previous approaches that employ a single equation can only use a single set of variables that account for confounding effects, implicitly assuming that the same set of factors affect both directions of a potentially endogenous relationship. In our model, a different set of factors can be selected according to relevant theory and/or previous empirical findings.

Without it being an identification statement, the model in Eqq. (1)-(4) is estimated using the following control variables. We account for other firm-specific effects using a set of control variables, $CV = (Growth, IntCov, Debt\ Ratio, Current\ Ratio, P/S\ , FA/TA\ , IA/TA)$, that stand for growth, interest coverage ratio, total debt ratio, current ratio, pricesales ratio, fixed assets over total assets, and intangible assets over total assets, respectively. Each equation has a unique set of control variables, captured by a_q, β_q, μ_q.

The first control variable is growth (Growth), which is an integral part of CFP, as higher growth (Easton 2004) is a measure of increased profitability. Smaller firms usually experience higher growth (Gupta 1969), while the link between CSR and growth depends on available investment opportunities (Branco and Rodrigues 2006). Furthermore, firms will prioritize financial stakeholder claims over social stakeholders (Artiach et al. 2010). Therefore, highly leveraged firms should be less likely to improve their CSR profiles, even though it would further decrease their tax liability. However, CSR might contribute to sustainability of earnings and therefore reduce the overall risk (Husted 2005). We control for capital structure (Debt Ratio) in all three variables and allowed risk in the form of interest payments (IntCov) to be a determinant of CSR. We also control for liquidity (Current Ratio) and the perception of the market regarding the quality of sales (P/S). Finally, management might prioritize real investments (Chung et al. 1998) or intangible investments (Branco and Rodrigues 2006), depending on their marginal contribution to market value (Mackey et al. 2007). We use fixed assets (FA/TA) to control for the impact of real investments on CFP and SIZE, and Intangible Assets (IA/TA) as a determinant of CSR.

4 Empirical Results

4.1 Initial Observations

Table 1 presents the descriptive statistics for the full sample and three sub-samples. Firms in advanced economies, namely the EU and the US, invest more in intangible assets, while real investments are more important for firms in the sample denoted as 'Rest Of the World' (i.e., ROW), which mainly consists of less developed economies. The total investment in fixed assets is larger in ROW (0.3007) than in the EU and the US (0.2597 and 0.2918), while the investment in intangible assets is more important in developed economies (EU: 0.6690, US: 0.6086, and ROW: 0.5784). This is consistent with *CSR* ranking across these regions. EU firms score higher on average (0.4305), followed by the US (0.3378) and ROW (0.2728).

Table 1 also provides insight on how investments in intangible assets are related to *CFP* and risk. Advanced economies exhibit higher *ROE* and *ROA* (*ROE* is 0.0615 in ROW, 0.1527 in the EU, and 0.1871 in the US), mainly due to a long left tail in ROW (skewness = −3.60). The US and the EU exhibit higher *P/B* ratios (ROW: 1.7347, EU: 3.3685, and US: 3.7155), mainly due to a shorter right tail (skewness = 6.52 and kurtosis = 64.89 in ROW). The major difference between ROW and the other two might be related to a lower investment in intangible assets and might be due to higher risk in ROW. This is a first sign that *CSR* and *CFP* might be correlated, but no safe conclusions can be drawn on whether more profitable firms are more likely to invest in *CSR* or whether *CSR* investments contribute to profit stability and thus to reduced risk.

Focusing on this, the risk statistics confirm the differences between ROW and the other two groups. Moreover, ROW, the EU, and the US appear to be progressively less risky, and firms in ROW exhibit considerably lower total debt ratios (0.5868) than their counterparts in the other two groups (EU: 0.6761 and US: 0.6111), while also sustaining higher current ratios (ROW: 1.8169, EU: 1.7147, and US: 1.5748). This increased risk taking might be due to increased profitability and increased ability to cover interest payments (interest coverage ratio = 0.0796 in ROW, 0.3272 in the EU, and 0.4512 in the US) or might be due to lower market risk. Extending this idea, a considerable difference is observed between ROW and the other two groups; firms in advanced economies are more profitable, show greater investment in intangible assets, operate at higher risk levels, and exhibit higher firm values. Again, *CSR* and *CFP* seem to be linked, and this is an attribute shared by risk taking. However, it is yet not clear whether lower overall risk allows greater investments in intangible assets, which increases profitability, or whether higher profitability is a determinant of risk taking and investments in intangible assets or is potentially endogenous.

Firm size could probably provide deeper insight on the link between *CSR* and *CFP*, revealing size-related effects. Table 1 reveals that it is higher in the US (7.4318) than in EU (7.2175) and in ROW (7.2051), which is consistent with profitability and market value. Table 2 shows that size is highly correlated with investments in intangible assets (0.5261), *CSR* ranking (0.2829), borrowing levels (0.5047 with debt ratio), and profitability (-0.0716 with *ROE* and -0.2282 with *ROA*). This indicates that size might indirectly affect *CSR* and *CFP* or their link or that it might be endogenous. This is further investigated in Fig. 1 by measuring the average market value (*P/B*) and profitability (*ROE*) across firm size and *CSR*. We first investigate the impact of size on earnings (*ROE*, Panel A, Fig. 1) and on the market valuation of the firm's net assets (*P/B*, Panel B, Fig. 1) which indirectly accounts for the market valuation of intangible assets and of *CSR* investments. *CSR* and *ROE* exhibit a U-shaped relationship, but this is not constant across size. For smaller firms, lower and higher *CSR* is linked with higher *ROE*, while mediocre *CSR* exhibits the lowest. In contrast, larger firms exhibit higher *ROE* with improved *CSR*. In between, higher and lower *CSR* is consistently associated with higher *ROE*, but the cut-off point where better *CSR* ranking is translated into higher *ROE* decreases as size increases.

These differences become more apparent upon examination of the sub-samples. In ROW, *ROE* decreases with higher *CSR* in relatively smaller firms, while a higher *CSR* ranking has a clearly positive impact on *ROE* in relatively larger firms. In the EU, the worst-performing firms seem to be those of medium size and medium or low *CSR*. In the US, the asymmetric effect of *CSR* on returns is more evident in smaller firms but larger firms seem to also benefit, almost monotonically, from a higher *CSR*. Two major conclusions can be drawn from this analysis.

First, the link between *CSR* and *CFP* appears to indeed be asymmetric and non-monotonic but not for all firm sizes. It appears to follow a U-shaped pattern for smaller-sized firms while exhibiting a non-monotonically increasing pattern as size increases. Consequently, we suggest that the U-shaped link reported in previous studies is non-monotonic across size and is observed in the tails of the size distribution. Second, there is an overall U-shaped pattern that exhibits a decreasing significance across size, indicating that there is a notable size effect in the way *CSR* and *CFP* interact. The tipping point of the U-shaped link decreases with larger size up to a point where *CFP* increases monotonically with *CSR* in firms of relatively higher size. Therefore, the U-shaped link is mainly observed in small firms, while the effect of decreasing returns due to mediocre *CSR* performance (Barnett 2007; Barnett and Salomon 2012) becomes less significant in larger firms. These two points become more apparent when we focus on market values (i.e., *P/B*).

4.2 Parametric Analysis

The previous section highlights several implications derived from non-parametric analysis, which could be summarized in the following principal concerns. The *CSR-CFP* relationship appears to be asymmetric and potentially endogenous. We observe that the importance of this notable asymmetric relationship between *CSR* and *CFP* decreases across a third dimension, namely, firm size, to a monotonically increasing function, exhibiting strong regional variations. The tipping point, where the marginal impact of *CSR* on *CFP* turns positive, appears to change across different size levels. However, such a non-conditional analysis cannot identify this tipping point, while it cannot account for endogeneity either. We further pursue this task in this section, aiming at measuring the direct and indirect impact of size on the *CSR-CFP* link.

4.2.1 Size and asymmetry

One of the major attributes of our model is that it explicitly models a potentially asymmetric impact of *CSR* on *CFP* (*ROE*) and allows for an additional factor ($Size$) to have a direct impact on *ROE* and on how *ROE* is affected by *CSR*. The first column of the second panel of Table 3 presents the estimation results for Eq. (1).

ROE appears to be asymmetrically affected by *CSR*. Low (high) *CSR* has a diminishing (−0.0363) (increasing (0.0181)) impact on *ROE*. This is consistent with previous literature (Barnett 2007) that reports a U-shaped relationship and we expand on this by measuring the exact inflection point for every firm, as well as by investigating how it changes across a third dimension. An estimate of 0.4131 for γ_0 indicates that this is an a-priori threshold level of *CSR* performance (assuming 0 $Size$). Any level below this has a diminishing impact on *CFP* and it is the inflection point where the marginal impact of *CSR* turns positive. This threshold value, however, does not remain stable across different levels of $Size$. The estimate of γ_1 is negative (−0.0420) and shows that the inflection point is lower for larger firms. This is consistent with previous studies (Brammer and Millington 2008) arguing that larger firms can more easily capitalize on their *CSR* investments. Table 3 also shows that these empirical findings are robust to alternative estimation methods.

Fig. 2 graphically depicts these findings and presents the parametric estimate of *ROE* according to Eq. (1) for different levels of $Size$. *ROE* decreases in low *CSR* levels and only starts increasing after a threshold is exceeded. According to the estimates of Eq. (4), $Size$ has a diminishing impact on the threshold value, which is lower for larger firms. This can be observed by a shift to the left of the minimum point.

4.2.2 Size and endogeneity

Would always be profitable to intensify the *CSR* strategy? According to the findings above, this would depend on the unique identification of the tipping point, how it interacts with the additional dimension considered and inevitably the existence of potential endogeneity. The second and third columns of the second panel of Table 3 report the estimates of the parameters of Eqs. (2)-(3) for the full sample.

CSR and CFP seem to be endogenously related. An estimate of 0.0127 for ROE indicates that firms with higher profitability are more likely to invest in CSR and thus achieve higher CSR performance. Higher CSR can also boost ROE but only in larger firms. In fact, higher CSR investments in smaller firms are associated with lower profitability. In addition, $Size$ appears to be endogenous to both CSR and CFP. An estimate of -0.0449 associates larger firms with lower earnings (ROE). This correlation is strong, and higher-earning firms indeed appear to be smaller (ROE on $Size$ is -0.0636). At the same time, an estimate of 0.1014 shows that larger firms are more likely to invest in CSR and are associated with higher CSR performance. This in turn can lead to larger size (CSR on $Size$ is 0.6241), indicating that CSR and size are endogenous.

Collectively, this reveals a spiral and endogenous relationship. Larger and more profitable firms are more likely to invest in CSR, which is expected to further increase $Size$, which in turn determines whether increased CSR investments will enhance CFP. Looking at it from a corporate finance point of view, CSR investments might be a preferred strategy with respect to CFP for firms that expect a negative impact of $Size$ on CFP due to their larger $SIZE$. Increasing $Size$ would not be a sensible option since this would decrease CFP. However, increased investments in CSR would enhance $Size$ and then indirectly enhance CFP or at least would mitigate the negative direct impact of $Size$ on CFP. In contrast, small firms should aim at increasing $Size$ because this would reduce the threshold value, which in turn would increase the marginal contribution of CSR on CFP.

4.3 Robustness: Regional, Firm-specific Factors and Generalizability

The major contribution of our model is that it allows, the data to uniquely identify the inflection point for every firm. The firm-specific focus is by construction free from cross-sectional differences and has the potential to capture several inconsistencies, previously reported in the literature. However, the estimates refer to a full sample estimation and cannot indicate whether the model is sensitive enough to capture regional, market, or firm-specific differences and whether the identification of the threshold variable contributes to a better description of the relationship between *CSR* and *CFP*.

Would the results be consistent if the model were applied in different samples or if the additional dimension consisted of more or different control variables, other than firm size? In this section, we estimate the model in different samples/datasets to test whether the model is sensitive enough to provide estimates of the parameters that capture regional/market/data-stylized effects. In addition, we appreciate that the choice of one specific control variable or this specification of non-linearity might be restrictive. We re-estimate the model with a different specification of Eq. (4) that considers a linear/constant specification as well as that the third dimension might include more or different variables, such as Research and Development (R&D) expenses.

4.3.1 Measuring financial performance (*CFP*)

We test the robustness of our findings by considering another accounting (internal) measure of managerial performance, the Return on Assets (*ROA*), and an indirect measure that accounts for external (market) valuation of the firm's assets and consequently of the value of its equity capital (*P/B*), which yield comparable results.

ROA is expected to be highly correlated with *ROE*; yet, it might be driven by different fundamentals (or by the same fundamentals in a different way), such as the proportion of debt, taxation, and depreciation (investments in fixed assets, *FA/TA*, or intangible assets, *IA/TA*). The estimation results are presented in the first panel (columns 4-6) in Table 4. Firms with higher *ROA* are found to be more likely to invest more in *CSR* (the estimate of *ROA* on *CSR* is 0.0440), which in turn might result in a larger size (the estimate of *CSR* on *Size* is 0.6403). The direct impact of *Size* on *ROA* is rather diminishing (−0.0180), but it can be mitigated by higher (above the threshold of $0.3890 - 0.0324xSize$) *CSR* performance, which is found to increase *ROA* (the estimate of high *CSR* on *ROA* is 0.0244). In contrast, when *CSR* performance is below the threshold, *ROA* is even lower (the estimate of the low *CSR* on *ROA* is −0.0319). The threshold value seems to be inversely (−0.0324) linked to *Size*.

In addition, most studies use *ROE* to measure *CFP*, as opposed to a market-based performance measure, such as Tobin's *Q* and the *P/B* ratio. The reason for this is the marginal contribution of *CSR* investments on profits rather than on the cost of goodwill in terms of improved social image/brand name. In addition, it also better matches the backward-looking character of our proxy for *CSR*, namely CSP. However, we recognize that if there is a "true" *CSR*-*CFP* relationship, especially on a firm level, it should also be reflected on market values, as well as on other *CFP* measures too. Therefore, we test the robustness of our findings to a market-based (Price to Book ratio (*P/B*)) measure that focuses on valuation. The findings (first three columns of Table 4) are magnified. The long-term, cross-sectional mean of *P/B* is 4.0702 for the full sample and is consistent with the non-parametric estimate in Table 1. In accordance with previous findings, low *CSR* has a diminishing (-0.0460) impact on *P/B*, whereas high *CSR* significantly boosts (0.0279) market values. The threshold value is still comparable at 0.4108, falling by 0.0421 for every unit increase in *Size*. These findings suggest that both the asymmetric relationship and the diminishing effect of size are greater in magnitude when market values are considered, which highlights the importance of *CSR* for investors. Furthermore, firms with high *P/B* ratio are more likely to invest in *CSR* (*P/B* on *CSR* is 0.0241). This investment will lead to larger size (*CSR* on *Size* is 0.6208), which, like in previous (*ROE*) findings, will have a dual impact on market value; a negative direct (*Size* on *P/B* is −0.1012) and a positive indirect (lower *CSR* threshold) impact.

4.3.2 Regional factors

We test the adequacy of the model in capturing potentially differential degrees of endogeneity and asymmetry by re-estimating the model in different market environments (ROW, EU, US), which implicitly account for regional differences and market-stylized factors. The second panel (columns 7-15) of Table 4 presents the estimation results, which confirm our previous findings. For example, ROE in the EU is found to increase *CSR* performance (0.0123), which in turn leads to greater size (1.0303) and higher financial performance (0.0166), when *CSR* is greater than the threshold value ($0.4177 - 0.0422xSize$), and vice versa (low *CSR* is −0.0362).

However, the major difference lies in the different measurements of the degree of convexity between *CSR* and *CFP* in each market. The maximum threshold value is consistently lower in advanced economies (ROW: 0.4680, EU: 0.4177, US: 0.4150). $Size$ is also found to significantly improve the profitability of *CSR* investments by lowering the threshold value (γ_1 is ROW: −0.0417, EU: −0.0422, US: −0.0425) and is associated with higher asymmetries (greater absolute difference) in advanced economies (*CSR-low* vs *CSR-high* on *ROE* is ROW: −0.0103 (−2.66) vs 0.0120 (2.65), EU: −0.0362 (-4.18) vs 0.0166 (3.83), US: −0.0649 (-4.88) vs 0.0233 (5.55)). These observations highlight the flexibility of our model in capturing varying degrees of non-linearity and endogeneity.

4.3.3 Inter-temporal differences, market-stylized factors and data samples

We test the robustness of our findings to alternative *CSR* scores, as well as their evolution in time and in different market environments. Previous studies report the inability of the existing *CSR* measures to capture the breadth of *CSR* (Carroll et al. 2016), as well as their failure to converge (Chatterji et al. 2016). If our model captures the "real" *CSR-CFP* link, its conditional propositions should be "on average" consistent independently of the *CSR* measure used and it should be flexible enough to capture changes in the intensity of the interaction. We address this issue in two ways. First, we employ different composite *CSR* scores, provided by different data vendors, which apply different criteria, as well as different weighting factors. Second, we decompose these scores into their constituent parts in order to investigate the ability of our model to capture more specific elements of the *CSR-CFP* relationship.

For the first part, we extend our analysis with another two samples consisting of firms listed in two S&P indices with valid observations for financial data and KLD and ESG (Bloomberg) scores for two periods; 1997-2010 (pre-2010) and 2011-2017 (post-2010). This approach tests the robustness of our findings to alternative market environments (e.g., the S&P indices), as well as potential evolution of the shape of the *CSR-CFP* relationship (e.g., between 1997-2010 and 2011-2017). Our model should be flexible enough to capture both.

The second panel of Table 5 presents the estimation results for the full model using the KLD, as well as the ESG rating, which confirm qualitatively the robustness of our results. We confirm the asymmetric impact of *CSR* on *CFP*, with the threshold value being a diminishing function of $Size$. The impact of a low *CSR* rating on *CFP* is again negative (*CSR*-low is −0.0272 with KLD and -0.0516 with ESG), while it increases (*CSR*-high is 0.0323 in KLD and 0.0385 in ESG) when *CSR* exceeds a threshold that is reversely related to $Size$ (γ_0 is 3.1504 and γ_1 is −0.6195 in KLD and 0.5050 and γ_1 is −0.0350 in ESG). The impact of $Size$ on the tipping point is plotted on Panels B and C of Fig. 2. Due to the range of the KLD measure the effect of the third dimension – shifting the threshold to the left – is even more pronounced, while it is highly similar (to Vigeo) when the ESG score is employed. However, the qualitative interpretation remains identical. This is a strong evidence that the relationship we report, using our model, is robust to measurement bias and temporal variation.

For the second part, we decompose each composite score into its constituent parts, as they are provided by the *CSR* rating agencies. Then, we replace in the estimations the composite score with each one of the sub-scores and report the results in Table 6, where the Vigeo score is decomposed into six domains: human resources (HR), environment (ENV), business behavior (C&S), corporate governance (CG), community involvement (CIN), and human rights (HRT). The KLD score is decomposed into strengths and concerns; each one computed as the sum of strengths or concerns, respectively. The ESG score is decomposed into three domains; Environmental, Social and Corporate Governance. The first panel (top half) of Table 6 presents the estimation results for the constituents of the Vigeo score, while the second panel (bottom half) presents the estimation results for the KLD and ESG scores.

The first notable observation is that our model appears to be flexible enough to capture differences on how each sub-score interacts, if at all, with *CFP*. This is observed when focusing on the estimates of the *CSR*-low vs *CSR*-high parameters in each one of the scores. Their magnitude and sign changes when a different *CSR* domain is employed. This shows that the model is flexible enough to capture both the shape and the intensity of the *CSR*-*CFP* link. For example, strengths (concerns) appear to monotonically increase (decrease) *CFP*. In addition, in accordance with Nollet et al. (2016), we find that the non-linearity is not persistent across sub-scores. In fact, we confirm that it is present when the corporate governance score is employed, but we report the same for the human resources score. In contrast, all other scores exhibit a monotonic relationship, which is persistent across different levels of magnitude of each score. This cannot be captured by pre-determined structural form, like the quadratic function used in previous studies (e.g., Barnet and Solomon 2012; Nollet et al. 2016) and it is a major merit of our modelling. Furthermore, a closer investigation of these differences reveals an additional feature that might be present in the *CSR*-*CFP* relationship. The two domains that exhibit a non-monotonic relationship, namely corporate governance and human resources, are internal to and can be controlled by the firm, while the other dimensions, e.g., the environmental and social interactions, that exhibit a monotonic link, lead to externalities. This provides some evidence that *CSR* actions that have externalities either increase or decrease *CFP*, according to the sign of the externality. In contrast, actions that are focused on the firm do not unconditionally lead to higher *CFP*. This might be somewhat intuitive, suggesting that actions that are easier (more difficult) to control, such as internal actions (external actions with externalities) might be more difficult (easier) to yield results, since their visibility/exposure is greater. This is a primary finding derived by using our model and we believe that it merits further investigation.

4.3.4 Functional Form

Finally, we test the robustness of our analysis with respect to how the additional dimension of the *CSR-CFP* nexus is modeled, in terms of functional form and control variables. We do this by considering a linear version of our model as well as extending Eq. (4) in two directions. First, we allow an additional, relevant variable, namely R&D expenses, to have an impact on the identification of the tipping point, as well as a direct impact on the endogenous variables. Second, we model this variable, assuming trifold endogeneity.

The estimation results of the linear specification are presented in the first 6 columns of Table 3, where $a_2' \equiv a_2$ and Eq. (4) is redundant. The estimation results remain qualitatively the same but the impact of *CSR* on *CFP* is found to be weaker (0.0079) and less significant (2.32), resulting in a worse fitting (Adj-R^2 and *MSE* are 20.58% and 0.1846 in the linear version and 40.12% and 0.1199 in the non-linear version). This is probably due to the fact that one parameter identifies a non-linear relationship (a_2 summarizes a_2'). Depending on the signs, the magnitude and the significance of the coefficients in a_2', a_2 might be positive, negative and/or insignificant, but most importantly sample dependent. Therefore, we suggest that the identification of the threshold level and consequently, the explicit design of the non-linearity might be the reason why previous (linear) studies might generate conflicting results. Our model captures sample dependent differences without sub-sample estimations and therefore it can capture a more fundamental link between *CSR* and *CFP*.

Naturally, the fitting of the model and its ability to adequately describe the *CSR-CFP* relationship depends on how the additional dimension is modeled. Towards this direction, we suggest an enhanced specification of Eq. (4) that accounts for another relevant control variable; R&D expenses. Recent literature investigates the marginal impact of R&D activity on *CFP* and *CSR* and considers R&D expenditures as investments in intangible assets (Chan, et al. 2001; Eberhart et al. 2004; Ehie and Olibe 2010), which contribute to differentiation and the development of competitive advantages and thus to longer term growth and profitability. We address this by including IA/TA as a control variable. However, more recent studies report that, when R&D investments are explicitly considered, the marginal impact of *CSR* on *CFP* becomes insignificant (McWilliams and Siegel 2000) or is only significant in low innovation firms (Hull and Rothenberg 2008), while R&D investments are positively correlated with *CSR* (Luo and Du 2015), exhibiting strong industry effects (Padgett and Galan 2010).

We test the above propositions by estimating an enhanced version of our model. There are two estimations. In the first, we add R&D as an explanatory variable in Eq. (1) and (2), as well as a determinant of s_{it} in Eq. (4). This extends our model by including more threshold variables and investigates how the *CSR*-*CFP* link changes across two dimensions. The second extension uses another dimension instead of *Size*, namely R&D investments, and aims at investigating whether the non-linear link still exists.

The empirical findings presented in the top panel of Table 7 suggest that, under the presence of R&D, which marginally increase *CSR* (0.0525 (2.34) in Vigeo, 0.5853 (4.66) in KLD and 0.2234 (3.01) in ESG), the impact of size is strengthened (0.1034 (7.92) in Vigeo, 0.1899 (6.70) in KLD and 0.1446 (3.56) in ESG). This suggests that larger firms invest more in *CSR*, either due to available resources (Chih et al. 2010) or because they have a lack of differentiation/innovation (Hull and Rothenberg 2008). With regards to the profitability of the *CSR* investments, they become profitable when *CSR* performance exceeds a threshold (0.0199 (4.96) in Vigeo, 0.0413 (2.92) in KLD and 0.0302 (3.77) in ESG). This threshold is negatively affected by *Size* (γ_1 is -0.0420 (-5.16) in Vigeo, -0.8943 (-5.50) in KLD and -0.0310 (-6.59) in

ESG) and now positively affected by the level of R&D investments (γ_2 is 0.0869 (2.90) in Vigeo, 0.4058 (4.34) in KLD and 0.0314 (3.94) in ESG). This means that larger firms still benefit more by *CSR*, but this effect is weaker. The threshold value is elevated upon the presence of significant R&D investments, which implies that firms with significant differentiation/innovation have alternative investment opportunities. These findings are also consistent when firm *Size* is replaced by R&D expenses (the bottom panel of Table 7) and highlight the flexibility of the model to capture various stylized factors.

5 Conclusions and Discussion

"Does it pay to be good?" This question refers to the marginal impact of *CSR* on *CFP*. Recent literature recognizes that *CSR* does not always contribute to higher *CFP* and therefore their link is asymmetric, usually U-shaped. Some studies attribute these asymmetries to endogeneity, while other studies consider that they are manifested by one or more interacting factors that can act as moderators/mediators. Although insightful, these studies operate under two restrictive conventions: (i) they consider a "universal" link between *CSR* and *CFP* that can be unconditionally applied to all firms, (ii) there is a trade-off between describing either the shape of the relationship or the interacting factors.

We address both issues by focusing on a more granular level: each firm. We introduce an additional dimension, which we proxy using an endogenous semi-latent specification, and we explicitly model how it affects the shape of the *CSR-CFP* relationship. This way we attain two contributions: (i) we suggest a flexible modelling that describes simultaneously the shape of the *CSR-CFP* link and the interacting factor, without imposing a functional form or making assumptions about what drives it. All inference is data driven. (ii) with our model we can identify explicitly, for the first time at a firm-specific level, the inflection point where the marginal impact of *CSR* on *CFP* turns positive. Consequently, any identified link would refer to a firm, accounting for firm specific factors, rather than an average cross-sectional depiction.

Our empirical findings confirm the asymmetric link between *CSR* and *CFP* but at a more granular level than previously thought. Indeed, there is a U-shape link between *CSR* and *CFP*, but its curvature is unique per firm. Therefore, the previously reported universally applied U-shape seems to be an "average" intensity curvature, which is the result of an aggregation of differently shaped individual relationships. This finding, which is robust to cross-sectional/measurement/specification differences, suggests that what previous studies identify "on average" is manifested because it holds at a firm level and when aggregated generates the previously observed U-shape. This firm level estimate is free from cross-sectional effects and therefore, it can explain previous conflicting results that arise due to these differences.

Finally, this, per firm, identification of the inflection point, which can also be conditional on any factors/key indicators of interest, has managerial implications too. Managers can identify key indicators that they believe will affect the profitability of their *CSR* investments. These indicators can be formulated into an additional dimension (i.e., an extended specification of Eq. (4)) and with this, they can identify the minimum threshold of *CSR* performance that they need to exceed in order for the marginal impact of *CSR* on *CFP* to be positive. With this, they can develop a profit maximizing strategy, relevant to the additional dimension. For example, according to our findings, smaller/higher R&D intensity firms might be better off if they first manage their Size/R&D intensity in order to reduce the *CSR* threshold, instead of simply intensify *CSR* investments.

References

Agle, B. R., Donaldson, T., Freeman, E. R., Jensen, M. C., Mitchell, R. K., & Wood, D. J. (2008). Dialogue: Toward superior stakeholder theory. *Business Ethics Quarterly*, 18, 153–190.

Aguilera, R., Rupp, D., Williams, C., & Ganapathi, J. (2007). Putting the s back in corporate social responsibility: A multilevel theory of social change in organizations. *Academy of Management Review*, 32, 836–863.

Aguinis, H., & Glavas, A. (2012). What we know and don't know about corporate social responsibility: A review and research agenda. *Journal of Management*, 38(4), 932–968.

Amato, L. H., & Amato, C. H. (2004). Firm size, strategic advantage, and profit rates in U.S. retailing. *Journal of Retailing and Consumer Services*, 11(3), 181–193.

Amato, L., & Wilder, R. P. (1985). The effects of firm size on profit rates in U.S. manufacturing. *Southern Economic Journal,* 52, 181–190.

Artiach, T., Lee, D., Nelson, D., & Walker, J. (2010). The determinants of corporate sustainability performance. *Accounting and Finance*, 50, 31–51.

Atkinson, L., & Galaskiewicz, J. (1988). Stock ownership and company contributions to charity. *Administrative Science Quarterly*, 33, 82–100.
Audretsch, D. B., Klomp, L., Santarelli, E., & Thurik, A. R. (2002). Gibrat's law: Are the services different? *Review of Industrial Organization*, 24(3), 301–324.
Aupperle, K. E., Carroll, A. B, & Hatfield, J. D. (1985). An empirical examination of the relationship between corporate social responsibility and profitability. *Academy of Management Journal*, 28, 446–463.
Avi-Yonah, R. (2005). The cyclical transformations of the corporate form: A historical perspective on corporate social responsibility. *Delaware Journal of Corporate Law*, 30(3), 767–818.
Barnett, M. (2007). Stakeholder influence capacity and the variability of financial returns to corporate social responsibility. *Academy of Management Review*, 32, 794–816.
Barnett, M., & Salomon, M. R. (2012). Does it pay to be really good? Addressing the shape of the relationship between social and financial performance. *Strategic Management Journal*, 33, 1304–1320.
Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99–120.
Bass, A. E., & Milosevic, I. (2018). The ethnographic method in CSR research: The role and importance of methodological fit. *Business & Society*, 57, 174–215.
Bénabou, R., & Tirole, J. (2010). Individual and corporate social responsibility. *Economica*, 77, 1-19.
Blau, P. (1970). Formal theory of differentiation in organizations. *American Sociological Review*, 35(2), 201–218.
Bloomberg (2020). https://www.bloomberg.com/impact/products/esg-data/. Accessed 14 January 2020
Bowen, F. E. (2002). Does size matter? Organizational slack and visibility as alternative explanations for environmental responsiveness. *Business & Society*, 41, 118–124.
Brammer, S., & Millington, A. (2005). Corporate reputation and philanthropy: an empirical analysis. *Journal of Business Ethics*, 61, 29–44.
Brammer, S., & Millington, A. (2008). Does it pay to be different? An analysis of the relationship between corporate social and financial performance. *Strategic Management Journal*, 29, 1325–1343.
Branco, M., & Rodrigues, L. (2006). Corporate social responsibility and resource-based perspectives. *Journal of Business Ethics*, 69, 111–132.
Capelle-Blancard, G., & Petit, A. (2017). The weighting of CSR dimensions: One size does not fit all. *Business & Society*, 56, 919–943.
Carroll, A. B. (1979). A three-dimensional conceptual model of corporate performance. *Academic Management Review*, 4, 497–505.
Carroll, R. J., Primo, D. M., & Richter, B. K. (2016). Using item response theory to improve measurement in strategic management research: An application to corporate social responsibility. *Strategic Management Journal*, 37, 66–85.
Chan, K., Lakonishok, J., & Sougiannis, T. (2001). The stock market valuation of research and development expenditure. *Journal of Finance*, 56, 2431–2456.
Chan, K. S., & Tong, H. (1986). On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, 7, 178–190.
Chatterji, A. K., Levine, D. I., & Toffel, M. W. (2009). How well do social ratings actually measure corporate social responsibility? *Journal of Economics & Management Strategy*, 18, 125–169.
Chatterji, A. K., Durand, R., Levine, D. I., & Touboul, S. (2016). Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Management Journal*, 37, 1597–1614.
Chih, H. L., Chih, H. H., & Chen, T. Y. (2010). On the determinants of corporate social responsibility: International evidence on the financial industry. *Journal of Business Ethics*, 93, 115–135.
Chung, K. E., Wright, P., & Charoenwong, C. (1998). Investment opportunities and market reaction to capital expenditure decisions. *Journal of Banking and Finance*, 22, 41–60.
Cornett, M. M., Erhemjamts, O., & Tehranian, H. (2016). Greed or good deeds: An examination of the relation between corporate social responsibility and the financial performance of U.S. commercial banks around the financial crisis. *Journal of Banking and Finance*, 70, 137–159.
Crane, A., Henriques, I., & Husted, B. W. (2018). Quants and Poets: Advancing Methods and Methodologies in Business and Society Research. *Business & Society*, 57, 3–25.
Darnall, N., Henriques, I., & Sadorsky, P. (2010). Adopting proactive environmental strategy: The influence of stakeholders and firm size. *Journal of Management Studies*, 47, 1072–1094.
Donaldson, L. (2001). *The contingency theory of organizations*. USA: Sage Publications.
Easton, P. (2004). PE ratios, PEG ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review*, 79, 73–95.
Eberhart, A., Maxwell, W., & Siddique, A. (2004). An examination of long-term abnormal stock returns and operating performance following R&D increases. *The Journal of Finance*, 59, 623–650.
Ehie, I. C., & Olibe, K. (2010). The effect of R&D investment on firm value: An examination of US manufacturing and service industries. *International Journal of Production Economics*, 128, 127–135.

Etzion, D. (2007). Research on organizations and the natural environment, 1992-present: A review. *Journal of Management*, 33(4), 637-664.
Evans, D. S. (1987). The relationship between firm growth, size and age: Estimates for 100 manufacturing industries. *Journal of Industrial Economics*, 35, 567–582.
Flammer, C. (2015). Does corporate social responsibility lead to superior financial performance? A regression discontinuity approach. *Management Science*, 61, 2549–2568.
Flammer, C, & Bansal, P. (2017). Does a long‐term orientation create value? Evidence from a regression discontinuity. *Strategic Management Journal*, 38, 1827–1847.
Fombrun, C, & Shanley, M. (1990). What's in a name? Reputation building and corporate strategy. *Academy of Management Journal*, 33, 233–258.
Freeman, R. E. (1984). *Strategic management. A stakeholder approach*. Marshfield, MA: Pitman.
Friedman, M. (1970). The social responsibility of business is to increase its profits. *New York Times Magazine*, September 13.
Galbreath, J., & Shum, P. (2012). Do customer satisfaction and reputation mediate the CSR-CFP link? Evidence from Australia. *Australian Journal of Management*, 37, 211–229.
Garcia-Castro, R., Ariño, M. A., & Canela, M. A. (2010). Does social performance really lead to financial performance? Accounting for endogeneity. *Journal of Business Ethics*, 92, 107–126.
García-Melón, M., Pérez-Gladish, B., & Gómez-Navarro, T. (2016). Assessing mutual funds' corporate social responsibility: a multistakeholder-AHP based methodology. *Annals of Operations Research*, 244(2),475–503.
Gibrat, R. (1931). *Les Inégalités Economiques*. Paris: Sirey.
Gully, M. S., Phillips, J. M., Castellano, W. G., Han, K., & Kim, A. (2013). A mediated moderation model of recruiting socially and environmentally responsible job applicants. *Personnel Psychology*, 66, 935–973.
Gupta, M. C. (1969). The effect of size, growth, and industry on the financial structure of manufacturing companies. *The Journal of Finance*, 24, 517–529.
Hillman, A. J., & Keim, G. D. (2001). Shareholder value, stakeholder management and social issues: What's the bottom line? *Strategic Management Journal*, 22(2), 125–139.
Hooghiemstra, R. (2000). Corporate communication and impression management – New perspectives why companies engage in corporate social reporting. *Journal of Business Ethics*, 27(1), 55–68.
Hull, C. E, & Rothenberg, S. (2008). Firm performance: the interactions of corporate social performance with innovation and industry differentiation. *Strategic Management Journal*, 29, 781–789.
Husted, B. W. (2005). Risk management, real options, corporate social responsibility. *Journal of Business Ethics*, 60(2), 175-183.
Isaksson, L. E., & Woodside, A. G. (2016). Modeling firm heterogeneity in corporate social performance and financial performance. *Journal of Business Research*, 69, 3285–3314.
Jawahar, N., Satish, P. G., Gunasekaran, A., & Subramanian, N. (2015). An Optimization Model for Sustainability Program. *Annals of Operations Research*, 250(2), 389–425.
Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: managerial behaviour, agency costs, and ownership structure. *Journal of Financial Economics*, 3, 305–360.
Jones, T. M. (1995). Instrumental stakeholder theory: A synthesis of ethics and economics. *Academy of Management Review*, 20, 404–437.
Lamata, M. T., Liern, V., & Pérez-Gladish, B. (2018). Doing good by doing well: a MCDM framework for evaluating corporate social responsibility attractiveness. *Annals of Operations Research*, 267(1-2), 249–266.
Lewbel, A. (1998). Semiparametric latent variable model estimation with endogenous or mismeasured regressors. *Econometrica*, 66, 105–121.
Li, Y., Gong, M., Zhang, X., & Koh, L. (2018). The impact of environmental, social, and governance disclosure on firm value: The role of CEO power. *The British Accounting Review*, 50(1), 60-75.
Luo, X., & Du, S. (2015). Exploring the relationship between corporate social responsibility and firm innovation. *Marketing Letters*, 26(4), 703–714.
Mackey, A., Mackey, T. B., & Barney, J. B. (2007) Corporate social responsibility and firm performance: Investor preferences and corporate strategies. *Academy of Management Review*, 32, 817–835.
Mansfield, E. (1962). Entry, Gibrat's law, innovation and the growth of firms. *American Economic Review*, 52(5), 1023-1051.
Marshall, A. (1961). *Principles of Economics*. London: Macmillan & Co.
McWilliams, A., & Siegel, D. S. (2000). Corporate social responsibility and firm financial performance. *Strategic Management Journal*, 21, 602–609.
Meznar, M., & Nigh, D. (1995). Buffer or bridge? Environmental and organizational determinants of public affairs activities in American firms. *Academy of Management Journal*, 38(4), 975–966.
Navarro, P. (1988). Why do Corporations Give to Charity? *Journal of Business*, 61, 65–93.
Ness, K., & Mirza, A. (1991). Corporate social disclosure: a note on the test of agency theory. *The British Accounting Review*, 23(3), 211–217.

Nollet, J., Filis, G., & Mitrokostas, E. (2016). Corporate social responsibility and financial performance: A non-linear and disaggregated approach. *Economic Modelling*, 52, 400-407.
Oll, J., Hahn, R., Reimsbach, D., & Kotzian, P. (2018). Tackling complexity in business and society research: the methodological and thematic potential of factorial surveys. *Business & Society*, 57, 26–59.
Orlitzky, M. (2001). Does firm size confound the relationship between corporate social performance and firm financial performance? *Journal of Business Ethics*, 33(2), 167–180.
Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. *Organization Studies*, 24(3), 403–441.
Padgett, R., & Galan, J. I. (2010). The effect of R&D intensity on CSR. *Journal of Business Ethics,* 93, 407–418.
Pfeffer, J., & Salancik, G. R. (1978). *The External Control of Organizations*. Harper & Row: New York, NY.
Porter, M. E., & Kramer, M. R. (2006). Strategy & society: The link between competitive advantage and corporate social responsibility. *Harvard Business Review*, 84(12), 78–92.
Rappaport, A. (1998). *Creating shareholder value: A guide for managers and investors*. New York: The Free Press.
Ratchford, B. T., & Stoops, G. T. (1998). A model and measurement approach for studying retail productivity. *Journal of Retailing*, 64, 241–263.
RiskMetrics Group (2010). How to Use KLD STATS & ESG Ratings Definitions.
Russo, M. V., & Fouts, P. A. (1997). A resource-based perspective on corporate environmental performance and profitability. *Academy of Management Journal*, 40(3), 534–559.
Rowley, T., & Berman, S. (2000). A brand new brand of corporate social performance. *Business & Society*, 39, 397–418.
Saeidi, S. P., Sofian, S., Saeidi, P., Saeidi, S. P., & Saaeidi, S. A. (2015). How does corporate social responsibility contribute to firm financial performance? The mediating role of competitive advantage, reputation, and customer satisfaction. *Journal of Business Research*, 68, 341–350.
Schreck, P., & Raithel, S. (2018). Corporate social performance, firm size, and organizational visibility: Distinct and joint effects on voluntary sustainability reporting. *Business and Society*, 57, 742–778.
Shahzad, A. M, & Sharfman, M. P. (2017). Corporate social performance and financial performance: Sample-selection issues. *Business & Society*, 56, 889-918.
Shepherd, W. G. (1972). The elements of market structure. *Review of Economics and Statistics*, 54, 25–35.
Solomon, R., & Hanson, K. (1985). *It's good business*. New York: Atheneum.
Spence, L. J. (2016). Small business social responsibility: Expanding core CSR theory. *Business & Society*, 55(1), 23-55.
Stigler, G. J. (1958). Economies of scale. *The Journal of Law & Economics*, 1, 54–71.
Surroca, J., Tribo, J., & Waddock, S. (2010). Corporate responsibility and financial performance: The role of intangible resources. *Strategic Management Journal*, 31(5), 463–490.
Trudel, R., & Cotte, J. (2009). Does it pay to be good? *MIT Sloan Management Review*, 50, 61–68.
Udayasankar, K. (2008). Corporate social responsibility and firm size. *Journal of Business Ethics*, 83(2), 167–175.
Van, D. D., Teräsvirta, T., & Franses, P. H. (2002). Smooth transition autoregressive models. A survey of recent developments. *Econometric Reviews*, 21,1–47.
Vigeo Eiris (2020). http://vigeo-eiris.com/. Accessed 14 January 2020.
Waddock, S. A., & Graves, S. B. (1997). The corporate social performance-financial performance link. *Strategic Management Journal*, 18, 303–319.
Walsh, G., & Beatty, S. (2007). Customer-based corporate reputation of a service firm: Scale development and validation. *Journal of the Academy of Marketing Science,* 36, 127–143.
Watts, R., & Zimmerman, J. (1986). *Positive accounting theory*. Englewood Cliffs, NJ: Prentice Hall.
Williamson, O. E. (1967). Hierarchical control and optimum firm size. *The Journal of Political Economy*, 75, 123–138.

Table 1 Descriptive Statistics (Vigeo)

		ROE	*CSR*	*Size*	*Growth*	*IntCov*	*Debt Ratio*	*Current Ratio*	*P/S*	*FA/TA*	*IA/TA*	*P/B*	*ROA*	*P/E*
Full Sample #7032	Mean	0.1459	0.3878	7.2593	0.0460	0.3151	0.6494	1.7017	1.8839	0.2724	0.6430	3.1924	0.0473	24.73
	Median	0.1272	0.3900	7.1795	0.0350	0.0548	0.6558	1.4439	1.1538	0.2033	0.6521	2.0155	0.0394	16.54
	Maximum	13.4428	0.7850	9.5040	1.1313	62.34	0.9967	49.9424	78.4050	0.9982	0.9987	239.13	1.1596	639.04
	Minimum	-6.1405	0.0533	4.8840	0.0000	-5.3613	0.0655	0.0079	0.0158	0.0001	0.0025	0.0650	-0.8905	0.4662
	Std. Dev.	0.4305	0.1286	0.7429	0.0507	2.2649	0.2007	1.5888	2.8304	0.2447	0.2124	6.6695	0.0745	39.26
	Skewness	11.64	-0.06	0.49	4.86	18.73	-0.31	13.87	9.74	0.93	-0.39	19.82	-0.23	8.27
	Kurtosis	342.74	2.34	3.23	61.33	419.79	2.56	307.34	186.03	3.00	2.64	561.56	28.12	92.41
ROW #1062	Mean	0.0615	0.2728	7.2051	0.0458	0.0796	0.5868	1.8169	1.8860	0.3007	0.5784	1.7347	0.0294	30.63
	Median	0.0641	0.2650	7.1254	0.0371	0.0027	0.5979	1.4819	0.8665	0.2577	0.5915	1.1930	0.0227	17.80
	Maximum	0.9239	0.6517	9.4125	1.1313	28.34	0.9895	49.9424	78.4050	0.9836	0.9925	30.73	0.3156	639.04
	Minimum	-1.7967	0.0533	5.7082	0.0000	-2.0078	0.0679	0.0545	0.0304	0.0005	0.0059	0.3022	-0.7035	1.48
	Std. Dev.	0.1715	0.1075	0.6110	0.0545	0.9327	0.2310	2.1149	3.7882	0.2364	0.2393	2.0929	0.0647	57.37
	Skewness	-3.60	0.35	0.69	8.65	26.93	-0.15	13.99	10.02	0.74	-0.19	6.52	-1.78	6.73
	Kurtosis	40.10	2.72	3.59	154.26	800.47	2.10	279.99	169.33	2.81	2.15	64.89	26.23	55.77
EU #4538	Mean	0.1527	0.4305	7.2175	0.0457	0.3272	0.6761	1.7147	1.7253	0.2597	0.6690	3.3685	0.0469	24.24
	Median	0.1354	0.4440	7.1216	0.0354	0.0628	0.6756	1.4474	1.0436	0.1981	0.6730	2.1487	0.0384	16.28
	Maximum	13.4428	0.7850	9.5040	0.9048	62.34	0.9957	42.1458	45.2815	0.9982	1.0907	239.13	1.1596	617.00
	Minimum	-6.1405	0.0900	4.8840	0.0000	-2.6976	0.0655	0.0144	0.0158	0.0001	0.0025	0.0650	-0.8905	0.47
	Std. Dev.	0.4293	0.1219	0.7914	0.0490	2.2609	0.1850	1.5802	2.4102	0.2329	0.1982	6.8553	0.0749	37.81
	Skewness	11.86	-0.35	0.47	4.49	18.91	-0.29	13.16	6.48	0.97	-0.40	18.54	0.11	8.31
	Kurtosis	355.48	2.57	3.02	47.71	430.99	2.67	265.25	77.37	3.25	2.78	498.66	32.04	94.05
US #1432	Mean	0.1871	0.3378	7.4318	0.0472	0.4512	0.6111	1.5748	2.3853	0.2918	0.6086	3.7155	0.0620	22.35
	Median	0.1471	0.3333	7.3548	0.0310	0.0731	0.6062	1.3983	1.6909	0.1719	0.6091	2.4252	0.0560	16.30
	Maximum	12.2944	0.6017	9.3599	0.4642	57.09	0.9967	23.4558	69.5716	0.9971	1.1813	227.30	0.7723	420.70
	Minimum	-5.8376	0.1183	6.1116	0.0000	-5.3613	0.0782	0.0079	0.0505	0.0008	0.0067	0.2664	-0.6502	0.53
	Std. Dev.	0.5451	0.0868	0.6420	0.0530	2.8796	0.2077	1.0692	3.1580	0.2816	0.2196	8.0224	0.0772	26.33
	Skewness	9.62	0.22	0.78	2.69	14.93	-0.15	7.88	12.30	0.86	-0.26	18.89	-0.72	6.60
	Kurtosis	217.22	2.64	3.60	14.27	260.61	2.34	137.65	240.80	2.41	2.48	467.00	20.01	66.82

Table 1 presents the descriptive statistics of all variables used in the estimation of the empirical model. The first panel presents the statistics for the full sample, while the following panels present the statistics for three sub-samples, namely Rest Of the World (ROW), Europe (EU) and the United States (US). The first ten variables are defined in Section 3 of the paper. The last three variables, namely P/B, ROA and P/E stand for Price-to-Book ratio, Return on Assets ratio and Price-to-Earnings ratio respectively.

Table 2 Correlation Matrix (Vigeo)

	ROE	*CSR*	*Size*	*Growth*	*IntCov*	*Debt Ratio*	*Current Ratio*	*P/S*	*FA/TA*	*IA/TA*	*P/B*	*ROA*	*P/E*
ROE	1												
	(---)												
CSR	0.0220	1											
	(1.84)	(---)											
Size	-0.0716	0.2829	1										
	(-6.02)	(24.73)	(---)										
Growth	0.0243	0.0486	-0.1885	1									
	(2.03)	(4.08)	(-16.09)	(---)									
IntCov	0.0321	-0.0155	-0.0887	0.0345	1								
	(2.69)	(-1.30)	(-7.47)	(2.89)	(---)								
Debt Ratio	0.0442	0.1754	0.5047	-0.2194	-0.1450	1							
	(3.71)	(14.94)	(49.02)	(-18.86)	(-12.29)	(---)							
Current Ratio	-0.0318	-0.0830	-0.0905	-0.1631	-0.0224	-0.1983	1						
	(-2.66)	(-6.98)	(-7.62)	(-13.86)	(-1.88)	(-16.96)	(---)						
P/S	0.0383	-0.0928	-0.1767	0.0601	0.0989	-0.2836	0.1362	1					
	(3.22)	(-7.81)	(-15.05)	(5.04)	(8.33)	(-24.80)	(11.52)	(---)					
FA/TA	-0.0245	0.0156	-0.1435	0.5797	-0.0076	-0.2261	-0.2435	0.0799	1				
	(-2.05)	(1.31)	(-12.16)	(59.66)	(-0.64)	(-19.46)	(-21.05)	(6.72)	(---)				
IA/TA	0.0277	0.1705	0.5261	-0.2336	-0.1498	0.9429	-0.1657	-0.3015	-0.2170	1			
	(2.32)	(14.51)	(51.87)	(-20.15)	(-12.70)	(237.38)	(-14.08)	(-26.51)	(-18.64)	(---)			
P/B	0.7311	0.0228	-0.1729	0.0534	0.0399	0.0622	-0.0449	0.1708	-0.0347	0.0252	1		
	(89.85)	(1.91)	(-14.72)	(4.49)	(3.34)	(5.22)	(-3.77)	(14.54)	(-2.91)	(2.11)	(---)		
ROA	0.4894	0.0358	-0.2282	0.1064	0.1363	-0.3354	-0.0101	0.1711	0.0178	-0.3456	0.2059	1	
	(47.05)	(3.01)	(-19.65)	(8.97)	(11.54)	(-29.85)	(-0.85)	(14.56)	(1.49)	(-30.88)	(17.64)	(---)	
P/E	0.0050	0.0039	-0.0132	0.0001	0.0015	-0.0008	0.0019	0.0097	0.0021	-0.0076	0.0065	0.0094	1
	(0.42)	(0.33)	(-1.11)	(0.01)	(0.12)	(-0.07)	(0.16)	(0.82)	(0.18)	(-0.64)	(0.55)	(0.79)	(---)

Table 2 presents the correlation matrix of all the variables employed in the empirical model for the full sample. The values in () are *t*-statistics.

Table 3 Estimation Results: Linear and Non-Linear Specifications (Vigeo)

	Linear-GMM			SUR			Non-Linear-GMM			SUR		
	ROE	*CSR*	*Size*	*ROE*	*CSR*	*Size*	*ROE*	*CSR*	*Size*	*ROE*	*CSR*	*Size*
Interc	0.1337	0.3706	7.2979	0.6296	0.4599	0.3226	0.1593	0.3702	7.2963	0.6217	0.4599	5.7675
	(1.61)	(8.88)	(13.70)	(10.04)	(6.98)	(5.13)	(1.76)	(8.87)	(13.83)	(9.49)	(6.98)	(15.66)
ROE		0.0101	-0.0624		0.0180	-0.1826		0.0127	-0.0636		0.0180	-0.1825
(P/B, ROA)		(4.02)	(-12.39)		(5.40)	(-11.80)		(4.51)	(-12.42)		(5.39)	(-11.79)
CSR-low	0.0079		0.6223	0.0082		2.5062	-0.0363		0.6241	-0.0944		0.0251
(CSR)	(2.32)		(9.97)	(2.42)		(9.25)	(-4.48)		(10.01)	(3.54)		(49.25)
CSR-high							0.0181			0.0241		
							(4.86)			(2.11)		
γ_0							0.4131			0.3412		
							(4.09)			(3.59)		
γ_1							-0.0420			-0.0316		
							(-5.41)			(-4.49)		
Size	-0.0370	0.1002		-0.1074	0.1168		-0.0449	0.1014		-0.0969	0.1168	
	(-2.79)	(14.77)		(-11.80)	(19.08)		(-3.20)	(10.78)		(-9.77)	(19.08)	
Growth	0.2040	0.1633	-0.9368	0.2643	0.1898	-0.8237	0.2092	0.1636	-0.9366	0.2843	0.1898	-0.8238
	(2.95)	(5.75)	(-3.04)	(2.11)	(6.02)	(-5.08)	(2.85)	(5.76)	(-3.03)	(2.26)	(6.02)	(-5.08)
IntCov		0.0002			0.0010			0.0002			0.0010	
		(0.31)			(1.70)			(0.32)			(1.70)	
Debt Ratio	0.0679	0.0369	0.6988	0.3361	0.0614	0.7777	0.0681	0.0371	0.6980	0.3342	0.0614	0.7776
	(1.47)	(1.59)	(10.59)	(10.17)	(2.92)	(18.58)	(1.48)	(1.60)	(10.58)	(10.11)	(2.92)	(18.58)
Current Ratio	-0.0021	-0.0206	-0.0112	-0.0082	-0.0210	-0.0058	-0.0033	-0.0208	-0.0111	-0.0081	-0.0210	-0.0058
	(-0.57)	(-0.29)	(-1.78)	(-2.37)	(-0.22)	(-1.30)	(-0.87)	(-0.30)	(-1.76)	(-2.37)	(-0.22)	(-1.30)
P/S	0.0047	0.1001	-0.0445	0.0109	0.3217	-0.0301	0.0049	0.1009	-0.0444	0.0110	0.3217	-0.0301
	(1.81)	(1.48)	(-5.47)	(5.50)	(5.76)	(-11.70)	(1.93)	(1.49)	(-5.46)	(5.54)	(5.76)	(-11.70)
FA/TA	-0.0153		-0.1060	-0.1217		-0.0354	-0.0222		-0.1061	-0.1230		-0.0354
	(-0.41)		(-1.86)	(-4.20)		(-0.98)	(-0.57)		(1.86)	(-4.24)		(-0.98)
IA/TA		-0.0124			-0.0398			-0.0124			0.0398	
		(-0.57)			(-2.02)			(-0.57)			(2.02)	
Adj. R^2	0.2058	0.2218	0.6439	0.2228	0.2861	0.6053	0.4012	0.4183	0.6439	0.4242	0.2861	0.6053
J-stat		0.3499						0.7876				
p		(0.95)						(0.85)				
MSE	0.1846	0.1618	0.4200	0.1810	0.1511	0.3282	0.1199	0.1618	0.4200	0.1208	0.1511	0.3282

Table 3 presents the estimation results for the model presented in Eq. (1)-(4). The first panel (first 6 columns) reports the estimation results for the linear specification, while the second panel (last 6 columns) presents the results for the non-linear specification. The table presents the estimation results using the Seemingly Unrelated Regressions (SUR) and the iterative Generalized Method of Moments (GMM) methods, with country, year and industry fixed effects. *T*-stats are reported in brackets. Each panel is dissected into four sections. The first reports estimates for the intercept, i.e., α_0, β_0, μ_0, the second the estimates for the endogenous variables, i.e., *CSR*, *CFP* and *Size*, the third one reports the estimates for the set of control variables, i.e., *CV*, while the last one reports the Adjusted R^2, the *J*-statistic, *p* is the probability of the *J*-test and the Mean Squared Error (*MSE*).

Table 4 Estimation Results: Alternative Performance Measures and Regional Sub-samples (Vigeo)

	Price to Book			Return on Assets			US			EU			ROW		
	P/B	*CSR*	*Size*	*ROA*	*CSR*	*Size*	*ROE*	*CSR*	*Size*	*ROE*	*CSR*	*Size*	*ROE*	*CSR*	*Size*
Interc	4.0702	0.3687	7.3217	0.0612	0.3648	7.2069	0.1819	0.4103	7.6483	0.1606	0.3360	7.3576	0.0611	0.2414	6.7883
	(1.99)	(8.79)	(13.69)	(2.74)	(8.53)	(13.17)	(2.49)	(11.19)	(8.89)	(2.70)	(9.09)	(8.71)	(1.13)	(7.91)	(7.01)
ROE		0.0241	-0.0787		0.0440	-0.0457		0.0249	-0.0402		0.0123	-0.0793		0.0052	-0.0111
(P/B, ROA)		(4.86)	(-13.02)		(4.20)	(-13.33)		(5.00)	(-10.06)		(4.34)	(-13.05)		(2.51)	(-7.96)
CSR-low	-0.0460		0.6208	-0.0319		0.6403	-0.0649		0.8939	-0.0362		1.0303	-0.0103		0.1514
(CSR)	(3.30)		(10.14)	(-8.77)		(10.31)	(-4.88)		(6.02)	(-4.18)		(12.18)	(-2.66)		(2.10)
CSR-high	0.0279			0.0244			0.0233			0.0166			0.0120		
	(3.71)			(6.68)			(5.55)			(3.83)			(2.65)		
γ_0	0.4108			0.3890			0.4150			0.4177			0.4680		
	(4.00)			(3.22)			(3.93)			(4.07)			(4.37)		
γ_1	-0.0421			-0.0324			-0.0425			-0.0422			-0.0417		
	(5.47)			(6.32)			(-5.53)			(-5.04)			(4.72)		
Size	-0.1012	0.1224		-0.0180	0.1277		-0.0534	-0.0305		-0.0456	-0.0828		-0.0368	-0.1801	
	(1.24)	(10.86)		(-7.31)	(10.08)		(-3.44)	(-8.12)		(-2.65)	(-9.87)		(-2.16)	(-12.38)	
Growth	0.7686	0.1630	-0.8952	0.2455	0.1622	-0.9114	0.3427	0.1851	-1.1339	0.2362	0.1650	-1.4071	0.1907	0.1320	-0.2618
	(3.22)	(4.74)	(-3.08)	(2.67)	(4.72)	(-2.98)	(2.92)	(5.97)	(2.78)	(2.52)	(5.12)	(3.69)	(2.48)	(4.57)	(-1.56)
IntCov		0.0002			0.0002			0.0002			0.0002			0.0003	
		(0.28)			(0.33)			(0.10)			(0.40)			(0.68)	
Debt Ratio	1.3425	0.0343	0.7595	0.0517	0.0373	0.6029	0.0855	0.0485	0.1495	0.0518	0.0338	0.7994	0.0487	0.0204	0.8824
	(0.86)	(1.47)	(10.58)	(3.86)	(1.61)	(9.45)	(1.61)	(1.82)	(7.43)	(1.63)	(1.49)	(8.82)	(1.41)	(1.30)	(13.64)
Current Ratio	-0.0404	-0.0204	-0.0102	-0.0016	-0.0217	-0.0081	-0.0056	-0.0316	0.0407	-0.0027	-0.0419	-0.0067	-0.0120	0.0329	-0.0128
	(-0.47)	(-0.28)	(-1.75)	(-0.84)	(-0.36)	(-1.26)	(-0.41)	(-1.94)	(2.88)	(-0.61)	(-2.63)	(-0.69)	(-1.19)	(1.12)	(-3.10)
P/S	0.1310	0.1095	-0.0488	0.0031	0.0786	-0.0418	0.0035	-0.0284	-0.0300	0.0051	0.1636	-0.0699	0.0077	0.1611	-0.0785
	(1.97)	(1.38)	(-5.73)	(2.54)	(1.16)	(-4.97)	(1.29)	(-0.50)	(-2.61)	(1.62)	(1.55)	(-5.67)	(2.14)	(2.03)	(-5.91)
FA/TA	-0.8462		-0.1200	-0.0122		-0.1210	-0.0198		-0.1019	-0.0231		-0.0276	-0.0744		0.3279
	(-0.70)		(-2.16)	(-1.76)		(-2.12)	(-0.38)		(-2.10)	(-0.52)		(-0.39)	(-1.71)		(3.84)
IA/TA		-0.0109			-0.0123			-0.0607			-0.0618			-0.0261	
		(-0.50)			(-0.56)			(-0.19)			(-2.41)			(-0.41)	
Adj. R^2	0.4408	0.2228	0.6420	0.3574	0.2255	0.6444	0.4226	0.3694	0.5242	0.3923	0.3253	0.4757	0.4052	0.1982	0.6045
J-stat		2.9832			3.4246			0.01			2.20			1.36	
p		(0.49)			(0.33)			(1.00)			(0.53)			(0.72)	
MSE	0.0058	0.1612	0.4172	0.4385	0.1616	0.4306	0.0558	0.0607	0.2791	0.1843	0.1344	0.5565	0.2155	0.1987	0.6061

Table 4 presents the estimation results for full non-linear model, presented in Eq. (1)-(4), with two alternative financial performance measures (first 6 columns), namely Price/Book ratio (*P/B*) and Return On Assets (*ROA*), as well as for three sub-samples (last 9 columns), namely Rest Of the World (ROW), Europe (EU) and United States (US), for the full Vigeo sample. The estimation method is iterative GMM, with country, year and industry fixed effects. T-stats are reported in brackets. Each panel is dissected into four sections. The first reports estimates for the intercept, i.e., α_0, β_0, μ_0, the second the estimates for the endogenous variables, i.e., *CSR*, *CFP* and *Size*, the third reports the estimates for the set of control variables, i.e., *CV*, while the last one reports the Adjusted R^2, the J-statistic with p being the probability of the J-test.

Table 5 Descriptive Statistics and Estimation Results (KLD and ESG)

	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis
				KLD (pre-2010)			
ROE	0.1155	0.1232	29.8544	-22.5226	0.6457	0.97	739.30
CSR	-0.2633	0.0000	15.0000	-12.0000	2.4151	0.79	8.09
Size	6.3208	6.2701	9.3736	3.9093	0.7605	0.38	3.30
Growth	0.0463	0.0332	12.9622	-1.7817	0.1358	77.98	7428.84
IntCov	0.6346	0.0718	79.8711	-46.6225	3.7590	10.89	172.44
Debt Ratio	0.5706	0.5586	0.9432	0.0003	0.2626	0.93	2.67
Current Ratio	2.3279	1.7148	93.2115	0.0222	3.0948	14.17	319.74
P/S	2.2477	1.4274	343.1143	0.0195	4.3066	47.46	3595.92
FA/TA	0.2455	0.1816	0.9954	0.0006	0.2223	1.12	3.55
IA/TA	0.1929	0.1305	0.9862	0.0001	0.1890	1.20	3.92
P/B	3.2438	2.2900	235.2100	0.1600	5.0930	19.86	652.74
ROA	0.0451	0.0507	2.7724	-1.3097	0.1810	-0.97	61.70
P/E	40.3691	20.0599	1680.87	0.1631	272.3798	39.84	198.36
				ESG (post-2010)			
ROE	0.2448	0.1744	5.2788	-0.5552	0.3768	7.53	78.45
CSR	0.3197	0.3017	0.7727	0.0289	0.1457	0.47	2.22
Size	4.1460	4.1035	5.8357	2.6782	0.4766	0.37	2.79
Growth	0.2184	0.0642	27.7273	-5.6129	1.4062	12.84	212.82
IntCov	0.4366	0.0979	109.8408	-0.8716	4.1800	21.65	501.89
Debt Ratio	0.2830	0.2669	0.8247	0.0000	0.1372	0.42	2.90
Current Ratio	1.7923	1.5049	11.8818	0.2081	1.0950	2.51	14.83
P/S	2.2554	1.8073	17.0369	0.0548	1.8563	2.37	12.46
FA/TA	0.4622	0.3464	1.9353	0.0095	0.3412	0.93	3.09
IA/TA	0.3014	0.2798	0.8892	0.0000	0.2131	0.46	2.34
P/B	6.1476	3.1999	759.6177	0.2802	25.5349	23.23	622.44
ROA	0.0738	0.0656	0.4337	-0.0829	0.0494	1.43	8.75
P/E	33.6495	19.0048	8051.3579	3.0920	256.3212	26.78	763.03

	KLD (pre-2010)			*ESG (post-2010)*		
	ROE	*CSR*	*Size*	*ROE*	*CSR*	*Size*
Interc	0.1219	0.9764	5.7560	0.0972	0.2609	3.8232
	(2.60)	(5.78)	(16.69)	(3.07)	(4.66)	(14.60)
ROE		0.3361	-0.0316		0.3118	-0.0253
		(10.63)	(-3.22)		(10.62)	(-4.71)
CSR-low	-0.0272		0.0115	-0.0516		0.0149
(CSR)	(10.73)		(4.27)	(-15.73)		(4.88)
CSR-high	0.0323			0.0385		
	(3.41)			(5.18)		
γ_0	3.1504			0.5050		
	(2.42)			(4.01)		
γ_1	-0.6195			-0.0350		
	(-8.26)			(-5.30)		
Size	-0.0283	0.1192		-0.0187	0.1584	
	(-3.64)	(4.23)		(-2.23)	(3.39)	
Adj. R^2	0.3278	0.4690	0.5911	0.3837	0.2675	0.2649
J-stat (p)		1.0243			0.9847	
(p)		(0.79)			(0.80)	
MSE	0.4130	0.5511	0.5278	0.2605	0.7623	0.1530

Table 5 presents the descriptive statistics and the estimation results for the model in Eq. (1)-(4) for two S&P indices data sample collected for the period 1997-2010 (pre-2010; sub-sample of S&P1500) and 2010-2017 (post-2010, S&P500), complemented by the KLD and the ESG (Bloomberg) CSR rating, respectively. All financial variables have been collected and treated, exactly the same way as in the estimation results presented in Table 3. Estimation is conducted using the iterative GMM method with the same control variables, as well as with country, year and industry fixed effects. The values in brackets are *t*-stats. The estimation results are dissected into three sections. The first reports estimates for the intercept, i.e., α_0, β_0, μ_0, the second the estimates for the endogenous variables, i.e., *CSR*, *CFP* and *Size*, also plotted in Figure 1, while the last one reports the Adjusted R^2, the *J*-statistic, p is the probability of the *J*-test and the Mean Squared Error (*MSE*). The KLD score has been computed as $CSR_{it} = \sum_{m=1}^{n} KLD_{mit}$, where KLD_m is the score (-1 or +1) of category m that a company i is rated according to, from *KLD*, at time t. The ESG score (Bloomberg) has been computed as $CSR_{it} = ESG_{it}/100$.

Table 6 Estimation Results: CSR Sub-scores (Vigeo, KLD and ESG)

	ENV			HR			HRT			C&S			CIN			CG		
	ROE	*CSR*	*Size*	*ROE*	*CSR*	*Size*	*ROE*	*CSR*	*Size*	*ROE*	*CSR*	*Size*	*ROE*	*CSR*	*Size*	*ROE*	*CSR*	*Size*
Interc	0.2007	0.3962	7.3004	0.0826	0.5222	7.1625	0.0800	0.5052	7.1248	0.1675	0.3627	7.1797	0.1526	0.1054	7.2869	0.0863	0.5026	7.1683
	(2.89)	(2.88)	(14.98)	(0.83)	(17.45)	(12.57)	(0.80)	(17.45)	(12.26)	(1.97)	(4.35)	(12.28)	(1.23)	(2.10)	(13.61)	(1.76)	(7.32)	(12.11)
ROE		-0.1550	-0.0761		0.1845	-0.0237		0.1878	-0.0037		-0.1544	-0.0497		-0.0908	-0.0648		0.2011	-0.0202
		(-2.61)	(-12.33)		(4.32)	(-11.07)		(4.32)	(-1.07)		(-4.02)	(-6.15)		(-1.54)	(-11.91)		(4.47)	(-1.91)
CSR-low	-0.0418		0.0060	-0.0234		0.0088	0.0032		0.0088	0.0033		0.0081	0.0017		0.0060	-0.0143		0.0087
	(-1.07)		(11.07)	(-3.31)		(10.72)	(1.31)		(10.72)	(1.84)		(10.78)	(1.34)		(11.02)	(1.95)		(10.50)
CSR-high	-0.0225			0.0114			0.0414			0.0182			0.0162			0.0199		
	(-2.50)			(4.15)			(4.50)			(5.06)			(2.33)			(4.59)		
γ_0	0.5382			0.4187			0.3871			0.2158			0.2937			0.4704		
	(4.09)			(4.49)			(3.94)			(1.91)			(1.94)			(6.97)		
γ_1	-0.0409			-0.0322			-0.0393			-0.0252			-0.0266			-0.0422		
	(-7.96)			(-5.26)			(-4.60)			(-2.01)			(-1.94)			(-7.72)		
Size	-0.0477	0.1309		-0.0785	-0.1851		-0.0479	0.2185		-0.0189	0.0765		-0.0627	0.3030		-0.0496	-0.1840	
	(-3.05)	(7.12)		(-3.84)	(-4.95)		(-3.08)	(14.55)		(-1.96)	(3.23)		(4.21)	(16.92)		(-3.98)	(-4.93)	
Adj. R^2	0.2887	0.3692	0.6428	0.2803	0.3974	0.6174	0.2980	0.4017	0.6074	0.3164	0.3128	0.6218	0.3238	0.4105	0.6244	0.3040	0.4105	0.6175
J-stat (p)		4.0391 (0.26)			1.1390 (0.77)			1.9044 (0.59)			0.0085 (1.00)			0.1286 (0.99)			0.0011 (1.00)	
	Strengths			Concerns			Environmental			Social			Governance					
Interc	0.1047	2.3830	5.2226	0.1392	6.7373	6.1379	0.0421	0.2572	4.0077	0.0215	0.2495	3.7706	0.0842	0.5725	3.7055			
	(2.09)	(6.90)	(13.38)	(2.68)	(3.92)	(19.40)	(2.26)	(3.53)	(14.76)	(2.20)	(2.35)	(12.48)	(4.26)	(5.25)	(12.24)			
ROE		0.5364	-0.0374		-0.6876	-0.0260		0.2890	-0.0249		0.3905	-0.0248		0.3452	-0.0252			
		(16.05)	(-3.58)		(-12.70)	(-2.70)		(9.81)	(-4.71)		(15.00)	(-4.73)		(-10.99)	(-4.75)			
CSR-low	0.0315		0.0371	-0.0245		-0.0140	-0.0423		0.0182	-0.1542		0.0173	-0.1542		0.0108			
	(10.84)		(5.88)	(-10.24)		(-2.82)	(-6.84)		(6.02)	(-2.08)		(5.53)	(-22.08)		(4.62)			
CSR-high	0.0373			-0.0674			-0.0451			2.8071			0.0807					
	(9.03)			(-2.41)			(-5.00)			(10.50)			(10.50)					
γ_0	17.7563			18.2563			0.4953			0.5750			0.4475					
	(2.30)			(2.74)			(4.20)			(4.21)			(3.12)					
γ_1	-2.5047			-5.7880			-0.0739			-0.0539			-0.0447					
	(-7.23)			(-7.99)			(-4.81)			(-4.63)			(-7.41)					
Size	-0.0281	-0.1678		-0.0293	-0.1472		-0.0279	0.1317		27.9312	0.1672		27.9312	0.1764				
	(-3.15)	(6.21)		(-2.47)	(-8.88)		(-2.27)	(3.02)		(9.27)	(3.46)		(9.27)	(3.27)				
Adj. R^2	0.4190	0.6205	0.3998	0.2959	0.4282	0.2890	0.4021	0.3097	0.3178	0.3816	0.2559	0.2736	0.2913	0.3083	0.2670			
J-stat (p)		1.7968 (0.62)			1.6939 (0.64)			0.8607 (0.83)			0.8883 (0.83)			0.9815 (0.81)				

Table 6 presents the estimation results for the model in Eq. (1)-(4) considering the constituent parts of the composite CSR scores for the Vigeo (top panel; dissected into scores for environment (ENV), human resources (HR), human rights (HRT), business behavior (C&S), social interactions (CIN) and corporate governance (CG)), the KLD (first 6 columns of bottom panel; dissected into strengths and concerns) and the ESG (last 6 columns of bottom panel; dissected into Environmental, Social and Governance scores) datasets. Estimation is conducted using the iterative GMM method with the same control variables, as well as with country, year and industry fixed effects. The values in brackets are *t*-stats. The estimation results are dissected into three sections. The first reports estimates for the intercept, i.e., α_0, β_0, μ_0, the second the estimates for the endogenous variables, i.e., *CSR*, *CFP* and *Size*, while the last one reports the Adjusted R^2, the *J*-statistic, p is the probability of the *J*-test and the Mean Squared Error (*MSE*).

Table 7 Estimation Results: Alternative Specifications (Vigeo, KLD and ESG)

	Vigeo			KLD (pre-2010)			ESG (post-2010)		
	ROE	*CSR*	*Size*	*ROE*	*CSR*	*Size*	*ROE*	*CSR*	*Size*
Interc	0.1363	0.3716	7.2991	0.1263	1.6723	5.7560	0.0881	0.2734	3.8232
	(8.36)	(8.83)	(13.61)	(2.69)	(9.60)	(16.69)	(2.79)	(3.88)	(14.60)
ROE		0.0198	-0.0628		0.3521	-0.0316		0.2987	-0.0253
		(3.79)	(-12.40)		(11.22)	(-3.22)		(6.02)	(-4.71)
CSR-low	-0.0157		0.6236	-0.0274		0.0113	-0.0493		0.0149
(CSR)	(-2.53)		(9.95)	(-11.06)		(4.23)	(-13.26)		(4.88)
CSR-high	0.0199			0.0413			0.0302		
	(4.96)			(2.92)			(3.77)		
γ_0	0.4155			3.8484			0.6604		
	(4.06)			(4.90)			(6.04)		
γ_1	-0.0420			-0.8943			-0.0310		
	(-5.16)			(-5.50)			(-6.59)		
γ_2	0.0869			0.4058			0.0314		
	(2.90)			(4.34)			(3.94)		
Size	-0.0316	0.1034		-0.0288	0.1899		-0.0182	0.1446	
	(-2.38)	(7.92)		(-3.72)	(6.70)		(-2.35)	(3.56)	
R&D	0.4687	0.0525		0.1239	0.5853		0.0863	0.2234	
	(2.01)	(2.34)		(6.33)	(4.66)		(4.35)	(3.01)	
Adj. R^2	0.4428	0.4181	0.6440	0.4227	0.4697	0.5914	0.4527	0.2155	0.2646
MSE	0.0910	0.1617	0.4204	0.2301	0.5502	0.5271	0.2480	0.7923	0.1451
	ROE	*CSR*	*R&D*	*ROE*	*CSR*	*R&D*	*ROE*	*CSR*	*R&D*
Interc	0.1833	0.3474	0.0493	0.1629	1.5838	0.0206	0.0574	0.3982	0.0392
	(7.02)	(10.01)	(2.09)	(3.31)	(9.58)	(2.37)	(2.83)	(4.63)	(5.88)
ROE		0.0177	0.0366		0.3180	0.0683		0.2566	0.0395
		(3.88)	(2.85)		(10.40)	(3.82)		(6.96)	(3.41)
CSR-low	-0.0046		0.9516	-0.0286		0.0081	-0.4279		0.0081
(CSR)	(-2.15)		(7.25)	(-9.81)		(2.01)	(-13.19)		(1.83)
CSR-high	0.0109			0.0374			0.3353		
	(2.34)			(3.16)			(2.67)		
γ_0	0.2774			0.1016			0.5519		
	(4.99)			(3.47)			(8.91)		
γ_2	0.1402			0.2883			0.3151		
	(2.57)			(2.85)			(2.12)		
Size									
R&D	0.3955	0.0561		0.1204	0.5906		0.1653	0.2695	
	(1.99)	(2.39)		(4.49)	(4.47)		(3.28)	(4.93)	
Adj. R^2	0.3879	0.3810	0.3636	0.3394	0.4105	0.4511	0.4876	0.2878	0.2674
MSE	0.1405	0.1784	0.6409	0.4780	0.6152	0.2435	0.2216	0.7554	0.1915

Table 7 presents the estimation results of an extended version of our model for all datasets employed, namely Vigeo, KLD (pre-2010) and ESG (post-2010). In the top panel, *R&D* is introduced in Eq. (1), (2) and (4). Eq. (4) can be written as $s_{it} = \gamma_0 + \gamma_1 Size_{it} + \gamma_2 R\&D_{it}$. In the bottom panel, $\gamma_1 = 0$ and $R\&D_{it}$ is considered endogenous. All estimations are conducted on the full sample, with iterative GMM and include control variables, country, year and industry fixed effects. Each section of estimation results is dissected into two sections. The first reports the estimates for the intercept and the endogenous variables. The second reports the Adjusted R^2 and the Mean Squared Error (*MSE*).

Fig. 1 ROE, P/B, Size and CSR (Vigeo)

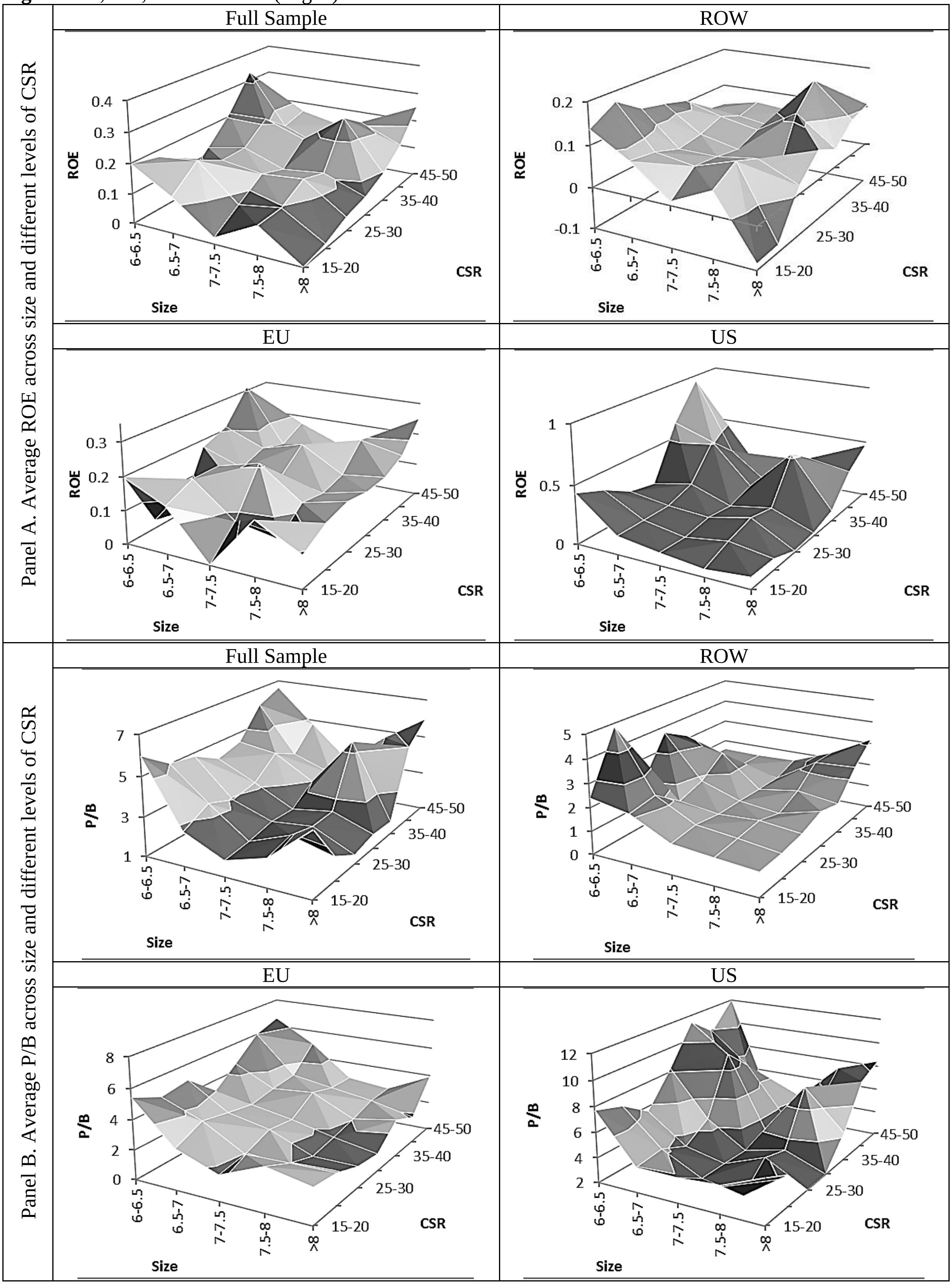

Fig. 1 presents the average ROE (Panel A) across different levels of Size and CSR for the full sample (Vigeo), as well as for the three sub-samples employed: Rest Of the World (ROW), Europe (EU) and the United States (US). Panel B presents the average P/B ratio across different levels of Size and CSR for the full sample, as well as for the three sub-samples employed: Rest Of the World (ROW), Europe (EU) and the United States (US).

Fig. 2 Return on Equity across different levels of CSR (Parametric) and firm Size (Vigeo, KLD and ESG)

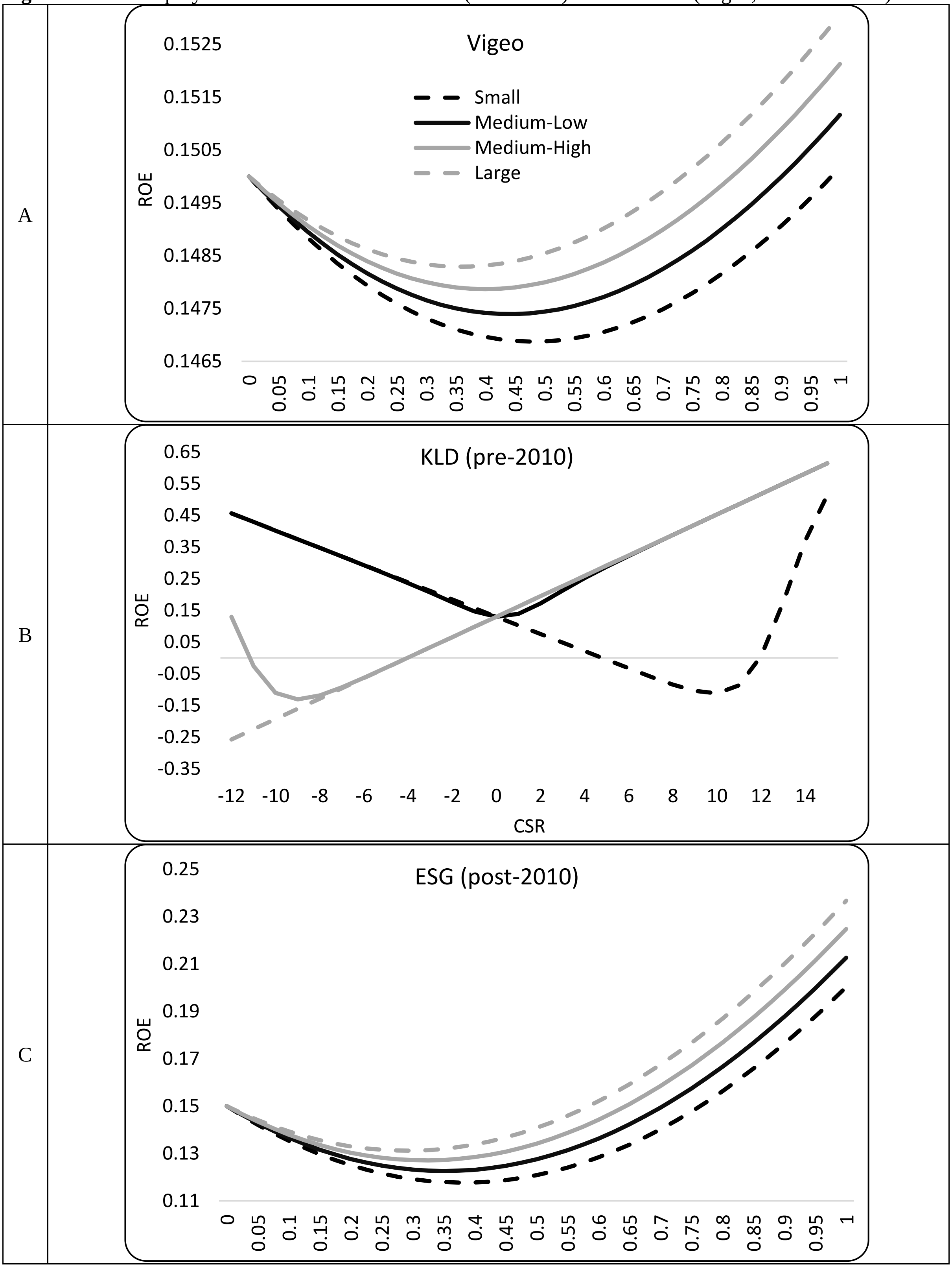

Fig. 2 presents the parametric estimate of *ROE* across different levels of *CSR* according to the estimates of the Equations (1)-(4), presented in Tables 3 and 5, for the Vigeo (Panel A), the KLD pre-2010 (Panel B) and ESG post-2010 (Panel C) datasets. *ROE* has been computed focusing on the marginal impact of *CSR* on *ROE* assuming an average α_0 equal to the unconditional average *ROE* and $a_{0,p} = a_q = 0 \; \forall \; p, q$. Firm size is dissected into Small ($Size = 3$), Medium-Low ($Size = 5$), Medium-High ($Size = 7$) and Large ($Size = 9$).

Technical Appendix
Online Supplement

The material presented in this document is intended to be used as an online supplement clarifying several aspects of quantitative nature, relevant to the construct and the estimation of our model.

Contents:

A. CSR composite scores

Vigeo

Vigeo Eiris is the leading European rating and research agency which evaluates companies based on Environmental, Social and Governance (ESG) criteria. Vigeo was established in 2002 and Eiris in 1983. Vigeo and EIRIS merged in 2015. In April 2019, Vigeo Eiris was acquired by Moody's Corporation, and Vigeo Eiris became a key part of Moody's ESG service (Vigeo Eiris 2020).

Vigeo measures CSR on a framework of 38 sustainability criteria, which are grouped into six broad domains of analysis: human resources (HR), environment (ENV), business behavior (C&S), corporate governance (CG), community involvement (CIN), and human rights (HRT). For each domain, there is a subset of criteria describing how the firm manages the particular aspect of CSR. Each criterion is activated and weighted according to its relevance by company sector. There are three factors which contribute to criterion weight: (1) nature of stakeholders' rights, interests and expectations; (2) vulnerability of stakeholders by sector; and (3) risk categories for the company. Vigeo's analysis focuses on how each company addresses each criterion in terms of leadership, implementation and results (Cavaco and Crifo 2014). The evaluation is realized by Vigeo via a questionnaire and not by the firms themselves. For each criterion, the questionnaire is structured into nine angles of analysis. The ratings model is based on internationally recognized CSR standards. (Eccles and Stroehle 2018). Vigeo provides six sub-scores and composite score. A score ranges from 0 to 100.

KLD

KLD methodology has been changed due to the transition to the MSCI ESG indices family that occurred on the 1st of September 2010; therefore in this paper, we collect KLD data until 2010 to maintain the consistency of KLD rankings among the sample companies from the S&P1500 index. For the details of the history of MSCI ESG indices, please see the discussion in Eccles and Stroehle (2018).

KLD dataset provides information about environmental, social and governance performance of companies. The KLD extended the coverage universe in 2003 to the 3000 largest US companies by market capitalization. KLD data were obtained from surveys, financial statements, media reports, regulatory filings, and other sources and were used to assess a firm's CSR performance on seven qualitative issues areas, namely community, corporate governance, diversity, employee relations, environment, human rights and product. Each of qualitative issues areas contains number of positive and negative indicators called strengths and concerns respectively. In addition, there is the eight category, namely controversial business issues which contains concerns only for six issues: alcohol, gambling, firearms, military, nuclear power, and tobacco. If a company did have strength or concern in a particular issue, then this is indicated with 1 otherwise 0. (RiskMetrics Group 2010).

Bloomberg ESG

Bloomberg evaluates companies on an annual basis, collecting public ESG information disclosed by companies through corporate social responsibility (CSR) or sustainability reports, annual reports and websites, and other public sources, as well as through company direct contact. This data is checked and standardized. Bloomberg ESG data covers 120 environmental, social and governance indicators including: carbon emissions, climate change effect, pollution, waste disposal, renewable energy, resource depletion, supply chain, political contributions, discrimination, diversity, community relations, human rights, cumulative voting, executive compensation, shareholders' rights, takeover defense, staggered boards, and independent directors (Bloomberg 2014). The weighting methodology is part of the proprietary calculation. It is weighted to emphasize the most commonly disclosed fields. The ESG score is based on 100 out of 219 raw data points collected. The weighted score is normalized to range from 0 (for companies that do not disclose ESG data) to 100 (for those companies which disclose every data point collected). Bloomberg accounts for industry-specific disclosures by normalizing the final score based only on selected set of fields applicable to the industry type. Bloomberg provides three sub-scores (Environmental, Social and Governance Disclosure Scores) and overall score (ESG Disclosure Score). Bloomberg ESG covers 11,000 companies in more than 100 countries (Bloomberg 2020).

B. The Model

$$\left\{\begin{aligned} CFP_{it} &= \left(\alpha_0 + a_{0,p}\sum I_{pit}\right) + \alpha_1 Z_{it} + a_2' CSR_{it} + \alpha_q \sum CV_{CFP,qit} + \varepsilon_{1it}, && (1)\\ CSR_{it} &= \left(\beta_0 + \beta_{0,p}\sum I_{pit}\right) + \beta_1 CFP_{it} + \beta_2 Z_{it} + \beta_q \sum CV_{CSR,qit} + \varepsilon_{2it}, && (2)\\ Z_{it} &= \left(\mu_0 + \mu_{0,p}\sum I_{pit}\right) + \mu_1 CFP_{it} + \mu_2 CSR_{it} + \mu_q \sum CV_{Z,qit} + \varepsilon_{3it}, && (3)\end{aligned}\right.$$

with

$$a_2' = a_2^{low}\left(\frac{1}{1+e^{\left(\frac{CSR_{it}-\{\gamma_0+\gamma_1 Z_{it}\}}{s_{it}}\right)}}\right) + a_2^{high}\left(1-\frac{1}{1+e^{\left(\frac{CSR_{it}-\{\gamma_0+\gamma_1 Z_{it}\}}{s_{it}}\right)}}\right) \quad (4)$$

C. Estimation

The model is estimated using an iterative Generalized Method of Moments (GMM) estimation method, as well as (linear and non-linear) Oridinary Least Squares (OLS) and Seemingly Unrelated Regression (SUR) methods, in order to test the robustness of our results. We define the moment conditions for the iterative-GMM estimations as follows. First, let $\boldsymbol{\beta} = (A, B, \Gamma, U)'$ be the vector of parameters to be estimated; $\boldsymbol{\varepsilon_w} = (\varepsilon_1, \varepsilon_2, \varepsilon_3)'$ be the vector of the error terms and $\boldsymbol{\pi} = (y_1, y_2, y_3)'$ be a vector of the fixed effects of each equation. $\boldsymbol{\upsilon} = (z_1, z_2, z_3)'$ is a vector of variables, z_w, orthogonal to the regressors of each equation $w = (1,2,3)$, computed as the OLS residuals of the regressions of each regressor on the dependent variable. These "artificial" instrumental variables are used as an additional level of treatment for endogeneity. The structural equation approach identifies the existences and magnitude of endogeneity, while the instrumental variables filter out time varying endogeneity of unknown form (e.g., Shahzad and Sharfman 2017), leaving the parameters to capture a "less biased" estimate of the interaction between *CSR* and *CFP*. In addition, let $G_{i,t}(s_{i,t}) \equiv \left(1 + \exp\left(CSR_{it} - \{\gamma_0 + \gamma_1 Z_{it}\}\right)\right)^{-1}$ and $I_{i,t} = 1_{CSR_{it}<\{\gamma_0+\gamma_1 Z_{it}\}}$, with $E\left(1_{CSR_{it}<\{\gamma_0+\gamma_1 Z_{it}\}}\right) = P(CSR < \{\gamma_0 + \gamma_1 E(Z)\})$ being the unconditional probability that *CSR* is below the threshold $s = \gamma_0 + \gamma_1 E(Z)$. The following moment conditions identify $\boldsymbol{\beta}$ and 3 constants C_w (constant variance).

$$E\begin{pmatrix} I_{i,t} - G_{i,t}(s_{i,t}) \\ \left(I_{i,t} - G_{i,t}(s_{i,t})\right)\boldsymbol{v_G} \\ \varepsilon_{w,it} \\ \varepsilon_{w,it} v_{w,it} \\ \varepsilon_{w,it}\pi_{w,i,t} \\ \varepsilon_{1,it}\varepsilon_{2,it}\varepsilon_{3,it} \\ \varepsilon_{w,it}^2 - C_w \end{pmatrix} = 0$$

The first two moment conditions, where $\boldsymbol{v_G}$ is a vector of orthogonal (to *CFP*) variables Z determining the threshold lelel $\{\gamma_0 + \gamma_1 E(Z)\}$, define the smooth transition function $G_{i,t}(s_{i,t})$ as the probability that CSR is below the threshold level (González et al. 2005) and identifies the parameter vector Γ. The third (set of) moment condition identifies the constants of each one of the three equations. The fourth (set of) moment condition identifies the vector of parameters A, B, M. The fifth moment condition imposes the restriction of no cross-correlation in the residuals, while the last (set of) moment condition assumes homoscedastic variances. In the case of the linear model, the first moment condition becomes redundant. The iterative-GMM is estimated with Newey–West heteroskedasticity–consistent errors. Hansen's (1982) J–statistics are used to test whether the moment conditions are well specified.

References

Bloomberg (2014). Bloomberg Look Beyond. https://cfaboston.org/docs/ESG/BloombergLookBeyond2014.pdf. Accessed 14 January 2020.

Bloomberg (2020). https://www.bloomberg.com/impact/products/esg-data/. Accessed 08 March 2020.

Cavaco, S., & Crifo, P. (2014). CSR and financial performance: complementarity between environmental, social and business behaviours. *Applied Economics*, 46(27), 3323-3338.

Eccles, R.G., & Stroehle, J. (2018). Exploring Social Origins in the Construction of ESG Measures. https://ssrn.com/abstract=3212685. Accessed 08 March 2020.

González, A., Teräsvirta, T., & Dijk, D. V. (2005). Panel smooth transition regression models. SSE/EFI Working Paper Series in Economics and Finance. http://swopec.hhs.se/hastef/papers/hastef0604.pdf. Accessed 31 January 2019.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 50, 1029–1054.

RiskMetrics Group (2010). How to Use KLD STATS & ESG Ratings Definitions.

Shahzad, A. M., & Sharfman, M. P. (2017). Corporate social performance and financial performance: Sample-selection issues. *Business & Society*, 56, 889–918.

Vigeo Eiris (2020). http://vigeo-eiris.com/. Accessed 08 March 2020.