

Hirsch, Stefan; Petersen, Thies; Koppenberg, Maximilian; Hartmann, Monika

Article — Published Version

CSR and firm profitability: Evidence from a meta-regression analysis

Journal of Economic Surveys

Provided in Cooperation with:

John Wiley & Sons

Suggested Citation: Hirsch, Stefan; Petersen, Thies; Koppenberg, Maximilian; Hartmann, Monika (2022) : CSR and firm profitability: Evidence from a meta-regression analysis, Journal of Economic Surveys, ISSN 1467-6419, Wiley, Hoboken, NJ, Vol. 37, Iss. 3, pp. 993-1032, <https://doi.org/10.1111/joes.12523>

This Version is available at:

<https://hdl.handle.net/10419/287984>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>

CSR and firm profitability: Evidence from a meta-regression analysis

Stefan Hirsch¹  | **Thies Petersen**¹  |
Maximilian Koppenberg¹  | **Monika Hartmann**² 

¹Department of Management in Agribusiness, Institute of Farm Management, University of Hohenheim, Stuttgart, Germany

²Chair of Agricultural and Food Market Research, Institute for Food and Resource Economics, University of Bonn, Bonn, Germany

Correspondence

Stefan Hirsch, Department of Management in Agribusiness, Institute of Farm Management, University of Hohenheim, Scherzstrasse 42, 70599 Stuttgart, Germany.
Email: s.hirsch@uni-hohenheim.de

Abstract

We analyze the literature on the relationship between corporate social performance (CSP) and corporate financial performance (CFP) by applying meta-regression analysis (MRA) to 7800 results of 512 empirical studies. Our findings reveal a small positive link between CSP and CFP as well as the presence of publication bias that favors statistically significant CSP–CFP regression coefficients. We also evaluate the impact of the underlying research design on the heterogeneity in published effects using Bayesian and frequentist model-averaging (FMA). We consider 42 contextual characteristics and our results show that reported CSP–CFP effects are smaller in cases where a binary index is used to measure CSP or when CSP is used as the dependent variable. In contrast, firms in industrial sectors or operating in China rather than the United States exhibit stronger effects. Finally, the CSP–CFP effect is driven by the choice of the econometric estimator and the inclusion of firm size as control variable.

KEYWORDS

corporate social responsibility, firm profit, meta-regression, publication bias

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Journal of Economic Surveys* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Corporate social responsibility (CSR) is a multidimensional concept referring to “the responsibility of enterprises for their impacts on society” (European Commission, 2011, p.6). CSR is recognized when co-operations in close collaboration with their stakeholders continuously commit to achieve a balance between their social, economic, and environmental responsibilities in an ethical manner (Dahlsrud, 2008; European Commission, 2011; Latapí Agudelo et al., 2019). CSR has been attracting considerable attention among business leaders, researchers as well as in the policy arena. Given an increased awareness for the significant impacts of companies on societies and the environment, stakeholders expect companies to live up to their CSR. At the same time, research findings indicate that key stakeholders such as consumers, employees, and investors reward firms’ responsible and punish their irresponsible conduct (e.g., Hartmann, 2011). This leads to the question whether there is a “business case” for responsible firm conduct.

In this article, we provide holistic evidence on the corporate social performance (CSP)–corporate financial performance (CFP) relationship by applying meta-regression analysis (MRA) to a sample of 7800 regression coefficients from 512 studies on the CSP–CFP relationship published between 1997 and 2022.¹ MRA is a regression approach that is particularly apt to provide a quantitative summary of the CSP–CFP regression coefficients and to identify the contextual factors that drive their heterogeneity (Stanley & Jarrell, 1989).² We derive 42 variables that capture various research design characteristics employed in the primary literature, such as the region and industry analyzed, the econometric estimator used, the type of CSP measure applied, or the control variables included, and assess their impact on the reported CSP–CFP effects.

From a theoretical perspective, there are three positions on the CSR–CFP relationship, the “traditional,” the “revisionist,” and the “slack resource” view (Lu & Taylor, 2016). The first involves neoclassical approaches, such as the theory of firms’ growth, which postulates that the superordinate goal of any firm is the generation of profits for its owners (Penrose, 1959). In line with the traditional view, Friedman sees “only one social responsibility of business—to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game, which is to say, engages in open and free competition without deception or fraud.” (Friedman, 1970, p.17). Based on agency theory (e.g., Ross, 1973), Friedman argues that managers are agents of the firm’s owner (principal) and have the obligation to serve the interests of the principal. With the engagement in CSR, managers follow their own goals. Hence, investments in social projects are considered as a reflection of an agency problem as this investment is assumed to lower the funds dedicated to operating the business thereby reducing corporate profits (Friedman, 1970). In contrast, the “revisionist” view suggests that CSR increases a firm’s positive image and reputation while decreasing transaction costs and thus, leads to an increase in profits. This view is based on legitimacy theory (Davis, 1973) and stakeholder theory (Freeman, 1984) with the former stating that firms seek to operate within the norms of their societal environment to secure their license to operate while the latter considers that a company acts in an environment where several stakeholders with varying needs exists—all relevant for the success of the company. According to the “revisionist” view, it is assumed that the better companies can integrate stakeholder needs into their business activities, the more successful they will be (Amoako & Dartey-Baah, 2020; van Beurden & Gössling, 2008). The resource-based view further supports the “revisionist” view as it suggests that close relationships with stakeholders enable a firm to develop specific intangible resources leading to competitive advantages and higher profitability (Wang et al., 2016). Finally, slack resource theory assumes a reversed causation between CSR and CFP by suggesting that it

is not CSR that drives CFP, but rather that firms with surplus financial means may spend it on CSR activities. In contrast, firms in financial distress are less likely to invest in CSR (Waddock & Graves, 1997; Wang et al., 2016).

Given the increasing relevance of CSR and the contradictory findings of the theoretical approaches an empirical research stream gradually evolved over the last four decades (e.g., Barnett & Salomon, 2012; Cavaco & Crifo, 2014) that investigates the relationship between CSP—the outcome of activities resulting from CSR (Carroll, 2018; Lu et al., 2014)³—and CFP to answer the question whether “it pays to be good.” Several review papers and meta-analyses have been conducted to summarize empirical findings on the CSP–CFP relationship concluding that the link is ambiguous as suggested by theory (Allouche & Laroche, 2005; Busch & Friede, 2018; Crifo & Forget, 2015; Dixon-Fowler et al., 2013; Endrikat, 2016; Friede et al., 2015; Grewatsch & Kleindienst, 2017; Hang et al., 2019; Hou et al., 2016; Huang, 2019; Huang et al., 2020; López-Arceiz et al., 2018; Lu & Taylor, 2016; Lu et al., 2014; Margolis et al., 2009; Orlitzky, 2001; Orlitzky et al., 2003; Vishwanathan et al., 2020; Wang et al., 2016). The majority of these meta-studies identify a small positive relationship between CSP and CFP (e.g., Friede et al., 2015). While most of the previous meta-analyses provide narrative summaries or descriptive analyses, only a few studies have used MRA to analyze the CSP–CFP relationship considering moderators. The existing MRA studies either focus on a specific region (e.g., Hou et al., 2016), are limited to one of the CSR dimensions (e.g., Hang et al., 2019), consider only few study design characteristics (e.g., Hou et al., 2016), or use a significantly smaller set of primary studies and observations as compared to our MRA (e.g., Allouche and Laroche, 2015). To the best of our knowledge, no previous MRA study on the relationship between CSP and CFP has been comprehensive in the inclusion of primary studies, moderators, and research design characteristics. An overview of previous MRAs on the CSP–CFP effect summarizing the core characteristics of these studies can be found in the appendix Table A1.

Our study identifies the link between CSP and CFP by conducting a MRA considering 42 variables potentially moderating the relationship between firms' social and financial performance. We account for model uncertainty by applying Bayesian and frequentist model-averaging (FMA). Our aim is to cover the full population of relevant empirical studies published up to 2022. By including 7800 regression coefficients from 512 primary studies, our analysis is far more comprehensive than all previous MRAs (c.f. Table A1) thereby increasing the reliability and robustness of the results on the CSP–CFP relationship. In contrast to earlier studies, we do not restrict our analysis to specific geographic regions or dimensions of CSP allowing us to identify the relevance of these factors for the link between CSP and CFP. Thus, our findings enable us to derive recommendations for firms' strategic orientation regarding investments in CSR activities. The insights generated are also of interest to policymakers seeking to promote voluntary CSR activities as well as evaluating the impact of politically enforced CSR measures on the competitiveness of firms. In addition, our MRA results facilitate the derivation of best-practice recommendations for future research on the CSP–CFP relationship. Finally, we assess the presence of publication bias, using recently developed nonlinear approaches. MRA can, therefore, be seen as a first step towards realizing a higher degree of transparency across the entire research stream by revealing potential misuse of statistical practices such as *p*-hacking. This is important, given the recent discussion on the abuse of *p*-values in the economic literature (Goodman, 2019).⁴

We proceed as follows. The next section describes the identification process for studies included in the MRA and the derivation of the dataset. This is followed by an analysis of publications bias in Section 3 and the investigation of heterogeneity in reported CSP–CFP effects in Section 4.

The article closes with some conclusions and implications for further research on the CSP–CFP relationship.

2 | LITERATURE SEARCH AND DATA

Relevant empirical studies on the CSP–CFP relationship were collected between January and April 2022 based on a systematic literature search. We followed the MRA guidelines formulated by Stanley et al. (2013) and Havránek et al. (2020) to collect all relevant published and unpublished (e.g., working papers and theses) articles that used a regression framework to empirically estimate the CFP–CSP relationship.⁵ We limit our study to regression analyses of the relationship as MRA can only be applied to a comparable measure that quantifies the same economic concept across primary literature (Oczkowski & Doucouliagos, 2015).

Our data collection focuses on the primary literature, which quantifies the CSP–CFP effect. The majority of primary studies employs the following regression model:

$$CFP_{k,t} = b_0 + b_{CSP}CSP_{k,t} + \sum_{m=1}^M b_m x_{m,k,t} + \varepsilon_{k,t} \quad (1)$$

where $CFP_{k,t}$ is the financial performance of firm k in period t . The r.h.s. of Equation (1) includes a single CSP variable and a vector x that reflects a set of $m = 1, \dots, M$ control variables that are assumed to have an impact on CFP.

Measurement of the CSP variable differs across literature since CSR is a multidimensional construct (Capelle-Blanchard & Petit, 2017; European Commission, 2003). Accordingly, the single CSP variable in Equation (1) can either be (i) one-dimensional capturing one of the three dimensions *environmental*, *social*, and *corporate governance*, (e.g., Capaldi et al., 2017), (ii) any of the three two-way combinations between the three dimensions, (e.g., Cavaco & Crifo, 2014), or (iii) an overall CSP score of the three dimensions (e.g., Feng et al., 2017).

In addition, some studies are based on Equation (2) and include two or three separate CSP variables that capture individual CSR dimensions allowing to disentangle the effect on CFP in a joint regression framework (e.g., Wang & Sarkis, 2017):

$$CFP_{k,t} = b_0 + \sum_{l=1}^3 b_{CSP_l}CSP_{l,k,t} + \sum_{m=1}^M b_m x_{m,k,t} + \varepsilon_{k,t} \quad (2)$$

where l reflects the three CSP dimensions.

Several studies also include interactions between CSP and specific control variables such as R&D expenditure to, for example, evaluate the relevance of CSR performance for the valuation of R&D expenditures (e.g., Ho et al., 2016):

$$CFP_{k,t} = b_0 + b_{CSP}CSP_{k,t} + \sum_{j \subseteq M} b_j x_{j,k,t} * CSP_{k,t} + \sum_{m=1}^M b_m x_{m,k,t} + \varepsilon_{k,t} \quad (3)$$

Note that Equation (3) can include one or several CSP dimensions. In our MRA, we focus on the main effect of CSP (\hat{b}_{CSP}) and additionally investigate the effect of including interaction terms on the main effect.

Finally, in some studies (e.g., Dupire & M'Zali, 2018), the effect is investigated based on the slack resource theory assuming reversed causality between CSP and CFP:

$$CSP_{k,t} = b_0 + b_{CSP}CFP_{k,t} + \sum_{m=1}^M b_m x_{m,k,t} + \varepsilon_{k,t} \quad (4)$$

In Equations (1)–(4), the estimated coefficient(s) \hat{b}_{CSP} show(s) the marginal link between CSP and CFP and is (are) hence the core parameter(s) of interest in our MRA.

In order to identify all studies that empirically analyze the above-indicated relationship based on Equation (1)–(4), the following databases were used for our literature search: Google Scholar, Econstor, SSRN, Jstor, Wiley, ScienceDirect, Web of Science, Springer, and Taylor and Francis. The complete set of relevant studies was identified using all reasonable combinations of the following key terms by combining them through Boolean strings using operators such as AND and OR: *CSR, CSP, (corporate) financial, performance, profit, and profitability*. We then checked the references in the resulting papers to detect other relevant studies using snowballing techniques (Longhi et al., 2005). This led to an initial set of 6506 potentially relevant articles. This set was reduced by excluding studies not covering the relationship between CSP and CFP and studies not written in English, as well as studies not presenting any kind of regression outcomes. Moreover, studies that used incomparable measures for CSP and CFP, studies that estimate Equation (4) as logistic regression with a binary CSP measure as dependent variable or studies with missing information on estimation precision (e.g., standard errors), were excluded since the availability of estimation precision is essential when testing for publication bias. We applied a multidimensional outlier detection algorithm (BACON) (Billor et al., 2000; Weber, 2010) to identify any unreasonable observations such as unrealistic coefficient-standard error combinations caused by reporting errors in the primary studies. This led to the exclusion of six individual coefficients, however, without excluding an entire study. The online Appendix describes the article inclusion process in more detail and provides a “Preferred Reporting Items for Systematic Reviews and Meta-Analyses” (PRISMA) template for the identification of included studies as well as descriptive information on the studies (e.g., Moher et al., 2009).

The majority of articles reports several CSP–CFP regression coefficients as their estimations are performed for different subsamples on various countries (e.g., Aouadi & Marsat, 2016), industries, different CSR dimensions, or with several econometric approaches in a single article. This leads to a final dataset consisting of 7800 CSP–CFP regression coefficients from 512 studies. Our data include articles published between 1997 and 2022 analyzing timeframes from 1990 to 2020.⁶ Hence, the literature search has yielded a considerable dataset that reflects the empirical CSP–CFP literature over the past three decades adequately. Note that 54.4% of the coefficients originate from estimations of Equations (1)–(3), that is, with CFP as a dependent variable, while 45.6% of estimates originate from a model where CSP is used as a dependent variable (Equation 4). Moreover, we have screened the titles and abstracts of the 512 articles to assess, if the study has an explicit focus on the CSP–CFP relationship since estimates from Equation (4) can also result from an analysis of the determinants of CSP other than profitability with the latter only being a control variable. We find that 58.1% of the included coefficients originate from studies where the main research focus is on the CSP–CFP relationship. We keep results from studies that have another main focus in the sample as they measure the same economic phenomenon and investigate potential differences in regression coefficients from these studies and those with a main focus on the CSP–CFP relationship.

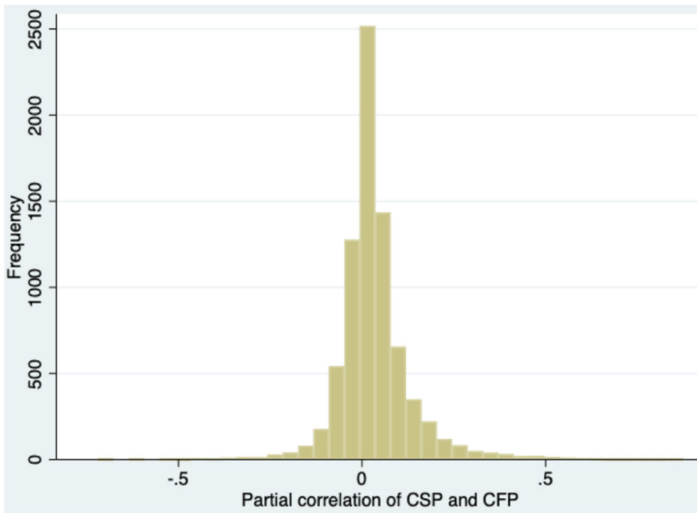


FIGURE 1 Frequency distribution for the partial correlation coefficient ($\hat{\beta}_{CSP}$): full sample. Note: $\hat{\beta}_{CSP}$ on the x-axis; the y-axis represents the frequency within the respective interval. $n = 7800$ [Colour figure can be viewed at wileyonlinelibrary.com]

The CSP–CFP regression coefficients reported in primary studies are used as the dependent variable in the MRA models. In the literature, a heterogeneous set of proxies with different measuring units was used to measure CFP and CSP (e.g., see Galant & Cadez, 2017). Therefore, the regression coefficients collected are not directly comparable in terms of magnitude. Since comparability of measured effects is a core requirement for a MRA (Havránek et al., 2020), we converted reported regression coefficients \hat{b}_{CSP} to partial correlation coefficients ($\hat{\beta}_{CSP}$) using (Greene, 2012, 37)

$$\hat{\beta}_{CSP} = \frac{t}{\sqrt{t^2 + df}} \quad (5)$$

where $t = \frac{\hat{b}_{CSP}}{SE}$. SE is the standard error of \hat{b}_{CSP} while df is the degrees of freedom of the underlying regression.⁷ In cases where z-statistics are used as a precision measure for \hat{b}_{CSP} (e.g., Barnett & Salomon, 2012), we follow Rosenthal (1991) and calculate $\hat{\beta}_{CSP}$ as follows:

$$\hat{\beta}_{CSP} = \frac{z}{\sqrt{n}} \quad (6)$$

where n is the sample size of the underlying regression. The resulting standardized effect $\hat{\beta}_{CSP}$ is distributed between -1 and 1 and represents the partial correlation between CSP and CFP (Wang & Shailer, 2015).⁸

Assessment of whether primary effects ($\hat{\beta}_{CSP}$) are affected by publication bias is based on the relationship to their estimation precision measured by standard errors, which are calculated by dividing $\hat{\beta}_{CSP}$ by its t -value (Stanley, 2005, 2008; Stanley & Doucouliagos, 2012).

Table A2 summarizes the partial correlation coefficients ($\hat{\beta}_{CSP}$) for the whole sample and separately for those cases where CSP relates to one specific CSR dimension. Mean $\hat{\beta}_{CSP}$ across all 7800 observations is 0.035 indicating a positive but small relationship between CSP and CFP. Figure 1 shows the frequency distribution of partial correlation coefficients revealing that they range from -0.72 to 0.88 (see also Table A2) centering around zero but with a positive skew.⁹ We also find that if the environmental and governance dimensions are considered alone (mean $\hat{\beta}_{CSP} = 0.024$ and

$\hat{\beta}_{CSP} = 0.023$), the $\hat{\beta}_{CSP}$ is significantly lower ($p < 0.01$) when compared to the social dimensions (0.035) (see Table A2).

In the case of 3.6% of estimated effects, only asterisks are provided to indicate the level of statistical significance, while no precise measures such as t -values, standard errors, or p -values are reported by the authors (see variable *test statistics* in Table A2). Therefore, we converted asterisks to t -values using the threshold they indicate. For example, if ** indicates $p < 0.05$, this was converted into the respective t -value for $p = 0.05$. This approach is somewhat inaccurate, since the exact p -value may have been any value between .01 and .05, which results in a higher t -value. Moreover, this also implies that only statistically significant results from the respective articles are included since a precision measure for insignificant findings cannot be derived in such cases. Therefore, we use a robustness check to assess the impact of this potentially inaccurate conversion on the MRA results.¹⁰

3 | PUBLICATION BIAS

Publication bias can originate from different sources, all of which can make the empirical effect of interest appear larger than it might actually be (Stanley, 2005). Journal editors and referees may possibly prefer to accept articles with statistically significant findings that, for example, are consistent with a specific theory (Card & Krueger, 1995). Therefore, publication bias can lead to a “file drawer problem” whereby authors tend to withhold nonsignificant results (Stanley, 2005).

3.1 | Descriptive analysis of publication bias

The presence of publication bias is indicated by a strong relationship between the estimates of interest, in our case, the partial correlations $\hat{\beta}_{CSP}$, and their estimation precision measured, for example, by standard errors (Stanley, 2005, 2008). A classical example illustrating this is t -statistics with an average absolute value around 2 (approximate threshold for statistical significance at the 5% level) across the literature of interest. Stanley (2005) suggests several graphical and numerical techniques to evaluate this relationship, whereby a funnel graph is the one most commonly used. A funnel graph plots the estimates against their estimation precision, which can be proxied by the reciprocal of the estimates' standard errors (Oczkowski & Doucouliagos, 2015; Stanley & Doucouliagos, 2010; Zigrainova & Havránek, 2016). If publication bias does not exist, estimated coefficients vary symmetrically and randomly around the value of the true population effect forming an inverted funnel. Estimates with high precision build the funnel's neck at the top of the graph and should be close to the true value. In turn, estimates with low precision will be more widespread at the bottom of the plot. The presence of publication bias will affect the symmetry of the graph by skewing the funnel in the direction of bias. On the other hand, if publication bias relates to statistically significant results regardless of the estimates' direction, the funnel will be hollow in the middle and widely spread over the entire range of $\hat{\beta}_{CSP}$ (Hirsch, 2018; Stanley, 2005).

The first step in the construction of a funnel graph involves deriving a proxy for the “true” population effect. In line with Stanley (2005), we average the top 10% of estimated values with respect to their precision to derive an indicator for the “true” CSP–CFP correlation. When considering the full sample, this average amounts to 0.023 indicating a small positive effect between CSP and CFP.

Figure 2 illustrates the funnel graph for the full sample of partial correlations ($\hat{\beta}_{CSP}$) whereby the vertical line at 0.023 depicts the approximation of the true value. The funnel is not skewed

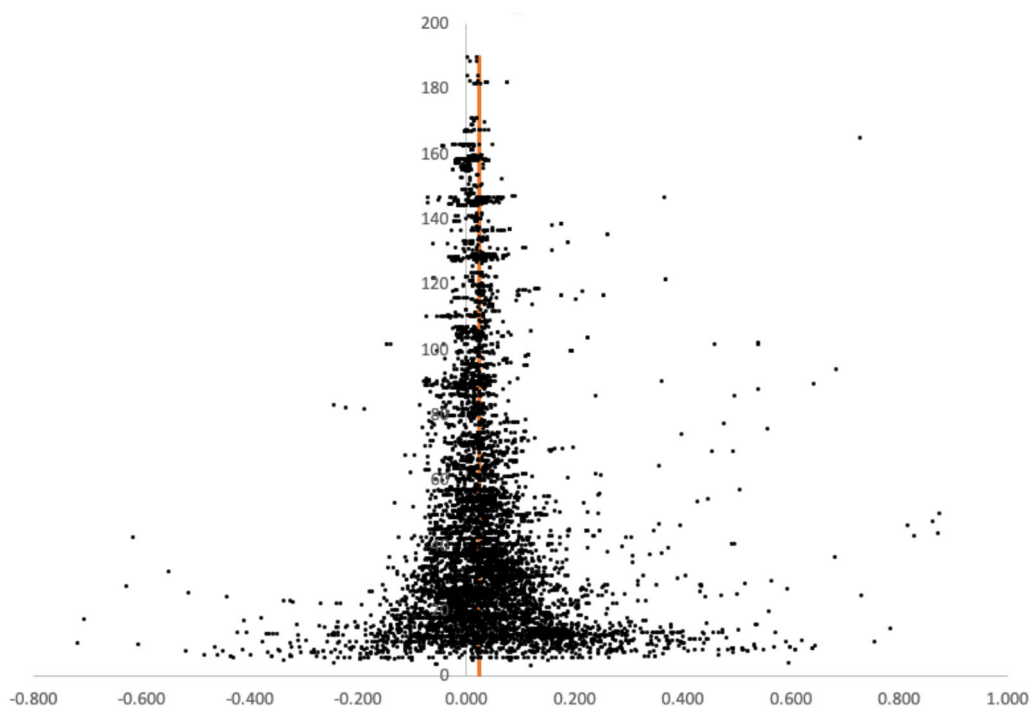


FIGURE 2 Funnel graph for the partial correlation coefficient ($\hat{\beta}_{CSP}$): full sample. Note: $\hat{\beta}_{CSP}$ on the x-axis; precision on the y-axis measured by the inverse standard error. The vertical line at 0.023 reflects the “true” $\hat{\beta}_{CSP}$ measured as the mean of the 10% most precisely estimated values. $n = 7800$ [Colour figure can be viewed at wileyonlinelibrary.com]

in either direction or hollow in the middle but appears widely spread over the entire range of estimates. Since both positive and negative partial correlations ($\hat{\beta}_{CSP}$) are supported by theory, and hence constitute “desirable” outcomes, this could be an indication of bi-directional publication bias in the empirical CSR-CFP literature.¹¹

The statistic $\left| \frac{\hat{\beta}_{CSP} - \beta_{CSP}}{se(\hat{\beta}_{CSP})} \right|$ represents a more formal way to detect an excessive likelihood of reporting significant $\hat{\beta}_{CSP}$ estimates, whereby β_{CSP} is the proxy for the “true” CSP–CFP correlation derived above by averaging the 10% most precisely estimated values. If there is no excessive likelihood of reporting significant $\hat{\beta}_{CSP}$ estimates, this statistic should not exceed 1.96 for more than 5% of the effects investigated (Stanley, 2005). The relevant percentages for the full sample and the individual CSP dimensions are summarized in Table A3. When viewing the full sample, it can be observed that the statistic exceeds 1.96 in 37.6% of the cases. Based on a z-test with H_0 that the proportion of values exceeding 1.96 (π) is equal to 0.05, this result is also statistically significant, which indicates an excessive tendency to report significant negative or positive $\hat{\beta}_{CSP}$ correlations. A similar finding can be observed for all subsamples of CSP dimensions where $H_0: \pi = 0.05$ is rejected in all cases.

In the next step, we apply several versions of the following MRA model (Stanley, 2005, 2008) to our data:

$$\hat{\beta}_{CSPi} = \beta_0 + \beta_1 se_i + \varepsilon_i \quad (7)$$

where the dependent variable is composed of the $i = 1, \dots, 7800$ partial CSP–CFP correlations ($\hat{\beta}_{CSPi}$) while ε_i reflects an i.i.d. error term. The $\hat{\beta}_{CSP}$ are regressed on their estimation precision measured, for example, by standard errors (se_i) to evaluate the presence of publication bias. If the latter is present, there will be a relationship between reported effects and their standard errors due to the fact that researchers seek out combinations of estimates and standard errors that yield a satisfactory level of statistical significance (Hirsch, 2018; Oczkowski & Doucouliagos, 2015). In Equation (7), this will be reflected by the rejection of $H_0: \hat{\beta}_1 = 0$. Testing H_0 is likewise an assessment of the funnel graph’s skewness and is also known as funnel-asymmetry test (FAT) (Stanley, 2005, 2008). This is based on the assumption that there is no correlation between $\hat{\beta}_{CSP}$ and se_i if publication bias was absent. However, the assumption might be violated by the definitions of $\hat{\beta}_{CSP}$ and se_i as well as a potential effect of methodological choices in the primary studies on $\hat{\beta}_{CSP}$ and se_i leading to spurious correlation between the two. Therefore, we additionally use the number of observations of the primary study as an instrument for se_i (Printezis et al., 2019; Stanley, 2005; Zigraiova et al., 2021). The constant $\hat{\beta}_0$ reflects the estimated “true” empirical CSP–CFP effect after controlling for publication bias. The test for $H_0: \hat{\beta}_0 = 0$ is, therefore, also known as precision-effect test (PET) (Stanley, 2005, 2008).

Further, Stanley and Doucouliagos (2012) suggest to use the squared standard error to capture a potential nonlinear relationship between estimates and standard errors:

$$\hat{\beta}_{CSPi} = \beta_0 + \beta_1 se_i^2 + \varepsilon_i \quad (8)$$

This specification leads to a more precise estimate of the “true” effect ($\hat{\beta}_0$) and allows publication bias to be less severe for more precise estimates (Hirsch, 2018; Kim et al., 2014).

The funnel graph indicates the presence of excess variation in reported CSP–CFP effects and hence the need to test separately for publication bias favoring positive and negative $\hat{\beta}_{CSP}$. Thus, to capture bi-directional publication bias, we extend model (7) by splitting the effect of se_i between the positive and negative domain of $\hat{\beta}_{CSP}$ using dummy variables $D_i^P = \begin{cases} 1 & \text{if } \hat{\beta}_{CSPi} \geq 0 \\ 0 & \text{if } \hat{\beta}_{CSPi} < 0 \end{cases}$ and $D_i^N =$

$\begin{cases} 1 & \text{if } \hat{\beta}_{CSPi} < 0 \\ 0 & \text{if } \hat{\beta}_{CSPi} \geq 0 \end{cases}$, respectively. This leads to

$$\hat{\beta}_{CSPi} = \beta_0 + \beta_2 D_i^P se_i + \beta_3 D_i^N se_i + \varepsilon_i \quad (9)$$

where the rejection of $H_0: \hat{\beta}_2 = 0$ indicates the presence of publication bias towards a positive CSP–CFP link and rejection of $H_0: \hat{\beta}_3 = 0$ towards a negative CSP–CFP relationship (Bom & Ligthart, 2009; Hirsch, 2018).

Two econometric problems must be addressed when estimating the above models. Since the dependent variable $\hat{\beta}_{CSPi}$ is composed of estimated effects with heterogeneous variances, the error terms are probably affected by heteroscedasticity leading to biased standard errors in the MRAs (Nelson & Kennedy, 2009). Given that se_i is a measure for this heteroscedasticity, corrected standard errors can be generated by applying weighted least squares (WLS) with $1/se_i^2$ as optimal weights (Hedges & Olkin, 1985; Oczkowski & Doucouliagos, 2015; Stanley, 2005). A second hurdle is that our sample is composed of clusters of partial correlations $\hat{\beta}_{CSPi}$ from 512 studies. Thus, cluster correlation is an additional source of bias for the standard errors (Nelson & Kennedy, 2009). Therefore, in line with Oczkowski and Doucouliagos (2015), we use study cluster robust and wild bootstrapped clustered standard errors, which are the adequate approach in cases where

TABLE 1 FAT-PET and nonlinear publication bias results

Panel A: Linear approaches						
	(1)	(2)	(3)	(4)	(5)	(6)
Variables	WLS SE	WLS SE ²	WLS No. of obs.	2SLS	WLS Bi-directional	Only CFP- dependent
Constant	0.017*** (0.004)	0.022*** (0.003)	0.021*** (0.003)	0.015*** (0.004)	0.013*** (0.004)	0.012***
se	0.490*** (0.167)			0.511*** (0.132)		0.646***
se ²		5.058*** (1.594)				
1/no. of obs.			6.083*** (1.628)			
se*D ($\hat{\beta}_{CSP} \geq 0$)					2.033*** (0.143)	
se*D ($\hat{\beta}_{CSP} < 0$)					-2.050*** (0.179)	
Observations	7800	7800	7800	7800	7800	4246
# of studies	512	512	512	512	512	295
Panel B: Nonlinear approaches						
	Andrews and Kasy	Stem-based	p-uniform	Endogenous- kink	WAAP	Average
Mean beyond bias	0.043***	0.009	0.045***	0.017***	0.019***	0.027
Standard error	(0.002)	(0.027)	(0.004)	(0.001)	(0.002)	
Observations	7800	7800	7800	7800	7800	
# of studies	512	512	512	512	512	

Note: Column (3) uses the inverse of the number of observations as direct instrument for se while column (4) uses 2SLS estimation with the inverse of the number of observations as instrument for se . For linear approaches, study-cluster robust standard errors are reported in parentheses. Andrews and Kasy (2019); stem-based (Furukawa, 2019); p -uniform (van Aert & van Assen, 2021); endogenous-kink (Bom & Rachinger, 2019); WAAP (Ioannidis et al., 2017). We also applied Bom and Rachinger's (2019) approach to publication bias on the negative side and find similar results.

*** $p < .01$, ** $p < .05$, * $p < .1$.

the number of clusters exceeds 42 (Angrist & Pischke, 2008). The wild bootstrap procedure is used as a robustness check since it is particularly well-suited to samples composed of clusters with a heterogeneous number of estimates, which is the case for our sample¹² (Cameron et al., 2008; MacKinnon & Webb, 2017).

Panel A of Table 1 provides results for the linear approaches to detect publication bias. In columns (1) and (2), the significant coefficients for se and se^2 reveal the presence of positive publication selection. This result is robust to using the inverse of the number of observations as a direct instrument for se_i and when the number of observations is used as an instrument for the standard error in a two-stage-least squares (2SLS) model. Moreover, column (5) indicates the existence of bi-directional publication bias confirming the excess variation revealed by the funnel plot (Figure 2). Finally in column (6), we only consider estimates that have been generated based on

Equation (1), that is, with CFP as the dependent variable leading to similar findings regarding publication bias. The constant ($\hat{\beta}_0$) ranges between 0.012 and 0.022 and is statistically different from zero in all models confirming the existence of a small positive CSP–CFP link after controlling for publication bias, which is smaller than the overall mean of reported estimates of 0.035 (see Table A2).

We also estimated Equation (7) separately for the subsamples of $\hat{\beta}_{CSP}$ relating to a specific CSP dimension. In each case, there is evidence for a small positive relationship, which is however insignificant for the case where $\hat{\beta}_{CSP}$ relates to the governance CSP dimension while publication bias is only detected for the social dimension (see Table A4). For the full sample, we also apply bootstrapped clustered standard errors to Equation (7) as a robustness check and observe that results do not change (see Table A4 column 4). Moreover, we exclude observations where only asterisks were provided to indicate the level of statistical significance of estimated effects to assess the impact resulting from this potentially inaccurate precision measurement. However, the results were not changed by the exclusion of the respective observations (see Table A4 column 5). Besides, we estimate Equation (7) only for observations from studies, which explicitly focus on the CSR–CFP relationship and for studies that are in addition estimating a model where CFP is the dependent variable (see Table A4 columns 6 and 7) leading to results resembling those for the full sample.

Finally, we apply five recently proposed nonlinear approaches that allow to derive average effects corrected for publication bias (e.g., Zigraiova et al., 2021). First, Andrews and Kasy (2019) build a selection model estimating the probability that insignificant results are underreported. Subsequently, this information is used to reweight insignificant coefficients. Second, Furukawa's (2019) stem-based method presents a novel approach to estimate the fraction of the top-10% most precisely estimated coefficients. Third, the p -uniform approach by van Aert and van Assen (2021) uses the distribution of p -values and searches for an underlying effect with a uniform distribution. Fourth, the “endogenous-kink” approach by Bom and Rachinger (2019) accounts for the fact that effects will only be reported if they exceed a specific precision threshold. It then estimates the effect under consideration that the relationship between estimates and standard errors is nonlinear below this precision threshold. Fifth, we apply the weighted average of adequately powered estimates (WAAP) approach by Ioannidis et al. (2017), which is based on the relationship of power and publication bias and a correction technique, which considers coefficients exceeding a power of 80%. Those coefficients are weighted by the reciprocal of their variance to arrive at a publication bias-corrected effect. In Panel B of Table 1, we report results for the mean effects beyond publication bias. The results vary between 0.009 for the stem-based method (Furukawa, 2019) and 0.043 for the Andrews and Kasy (2019) approach hence revealing a positive CSP–CFP effect after controlling for publication bias. The average across the five nonlinear methods of 0.027 is closer to the overall mean of reported estimates of 0.035 than the constant of the linear model in column (1) (0.017). This indicates that the degree of positive publication bias detected by the nonlinear approaches is somewhat lower than suggested by the linear approaches.

4 | HETEROGENEITY ANALYSIS

4.1 | Estimation and variables

Thus far, we have assumed a single underlying true effect. However, the “true” effect can differ across regions, industries, or time periods analyzed, thereby impeding the identification of the

“true” effect and publication bias within these subpopulations (Doucouliagos et al., 2005; Fernau & Hirsch, 2019; Hirsch, 2018; Stanley, 2005). In the following, we control for the underlying research design with a full MRA, which includes a vector of variables (k_j) related to the research design applied to estimate $\hat{\beta}_{CSP}$:

$$\hat{\beta}_{CSPi} = \beta_0 + \beta_1 se_i + \sum_{j=4}^n \beta_j k_{j,i} + \varepsilon_i \quad (10)$$

where the coefficient vector $\hat{\beta}_j$ captures the heterogeneity in $\hat{\beta}_{CSPi}$ caused by research design characteristics after controlling for publication bias. Finally, in Equation (10), the constant $\hat{\beta}_0$ reflects the “true” empirical CSP–CFP effect after controlling for publication bias and assuming all research design variables included in k_j are equal to zero (Doucouliagos & Stanley, 2009).

Since there is no theoretical foundation for the inclusion of independent variables in the MRA, we also apply Bayesian model-averaging (BMA) and FMA techniques as robustness checks.

According to Ullmann (1985) and Lu and Taylor (2016), heterogeneity in reported CSP–CFP coefficients is probably due to differences in the industries analyzed, varying definitions of CSP, omission of important variables as well as deficiencies in empirical databases and econometric methods. Therefore, we collected all those factors from the 512 studies included that best capture the research design characteristics employed and coded them as variables. They are illustrated in the underlying conceptual framework shown in Figure 3 and Table A2 in the appendix.¹³

The research design variables are grouped into four categories, whereby the first category includes dummy variables capturing general study characteristics related to article type and quality. While the majority of $\hat{\beta}_{CSP}$ originate from peer-reviewed journal articles (97.9%), 2.1% originate from non-peer-reviewed work (including working papers, book chapters, and theses) (cf. Table A2). Since it is reasonable to assume that working papers remain unpublished due to quality issues in the research design, the effect of publication quality can be evaluated by comparing these papers with those articles which appear in peer-reviewed journals (Havranek et al., 2018). Given that authors really intend to have their work published eventually in a peer-reviewed outlet, it is not possible to detect publication bias and the “file-drawer problem” (which arises when studies producing null results remain in the researcher’s cabinet) by differentiating between published and working papers (Fernau & Hirsch, 2019; Rusnak et al., 2013; Valickova et al., 2015).¹⁴

Previous meta-studies have failed to detect any relationship between reported estimates and journal quality measured by the number of citations (Hang et al., 2019) or impact factors (IFs) (e.g., Suñé et al., 2013). Nevertheless, we include a dummy variable that indicates for each $\hat{\beta}_{CSP}$ whether it originates from a journal with an IF at the time of publication. Table A2 indicates that 77.6% of $\hat{\beta}_{CSP}$ relate to such journals.

The second class of variables relates to the study focus covering the time frame, region, and industry analyzed. Brower and Dacin (2020) and Friede et al. (2015) state that the significant increase in the awareness of CSR activities and learning effects has led to a gradual weakening of the relationship with CFP over time. Hence, to assess changes in the link between CSP and CFP, we include the final year of the underlying sample for each $\hat{\beta}_{CSP}$. We subtract 1989 as the base year to simplify the coefficient estimate’s interpretation. Table A2 shows that the average final year is 2013 with a range from 1990 to 2020.

The relationship between CSP and CFP is likely to differ across industries (Baird et al., 2012) since companies operate in different economic contexts with diverse stakeholder groups across sectors (Baird et al., 2012; Feng et al., 2017). For example, CSR probably has a strong impact in

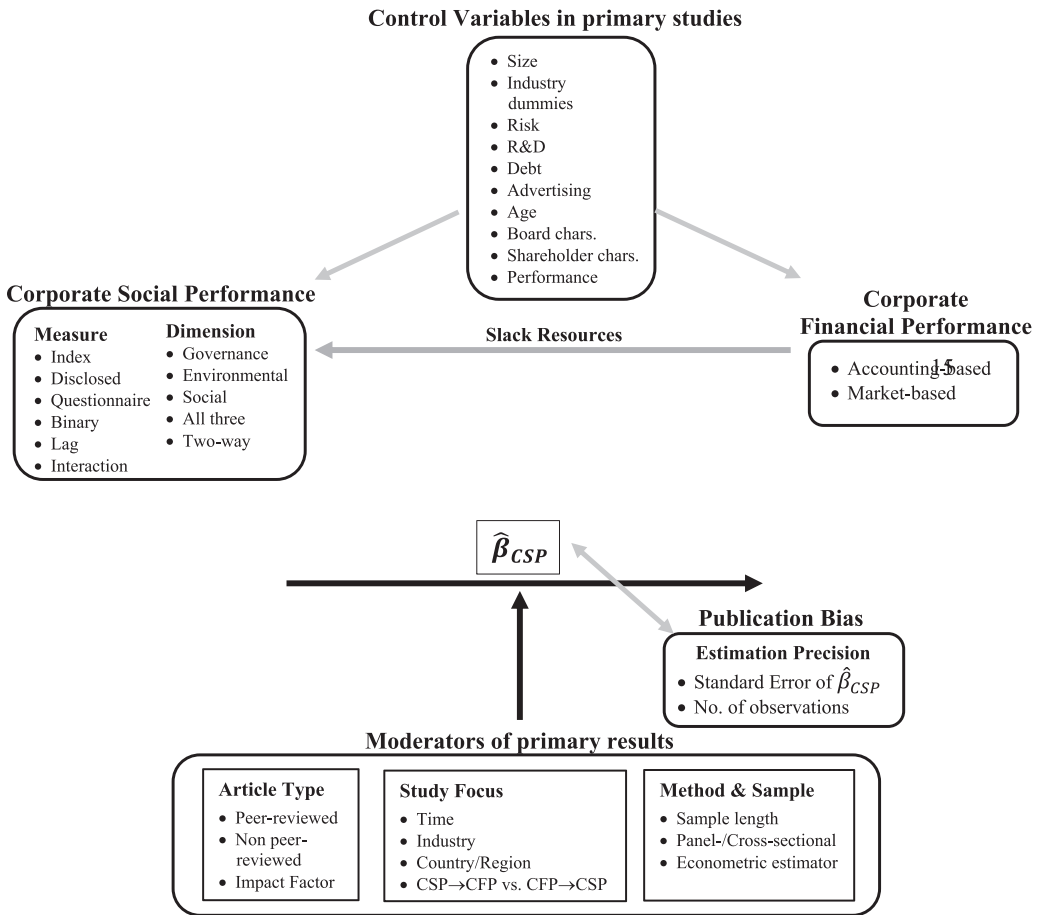


FIGURE 3 Meta-regression analysis framework

industrial sectors due to higher stakeholder attention (Hartmann, 2011; Hou et al., 2016). CSR is also a crucial factor in terms of consumers' perception and is thus a prime issue for the competitiveness of firms in the service sector (Musso & Rizzo, 2006). Moreover, CSR strategies and asset structures of financial firms differ from those of firms in other industries (Belu & Manescu, 2013). Therefore, we used the Global Industry Classification Standard (GICS) to account for industry-specific differences in the CSP–CFP relationship and grouped the effects based on the firms analyzed into industrials (GICS 20) (3.0%), financials (GICS 40) (8.2%), and consumer goods (GICS 25 and 30) (3.2%). The remaining effects (85.5%) are generated based on multi-industry samples that cannot be further disaggregated (e.g., Aouadi & Marsat, 2016), whereby multi-industry samples may obscure industry-specific effects and lead to a weaker relationship between CSP and CFP (Allouche & Laroche, 2005). Hence, assessing differences in coefficients between single and multi-industry studies allows us to evaluate biases in the measurement of the CSP–CFP effect using multi-industry samples.

The literature has also focused on various countries with different institutional settings (see the description of the included literature in the online appendix). Matten and Moon (2008) highlight that the adoption of CSR strategies differs strongly between firms in Europe and the United States. While the latter tend to adopt “explicit” CSR activities, EU companies are more engaged in

“implicit” CSR (Matten & Moon, 2008).¹⁵ Therefore, we include country dummy variables, which indicate that 14.8% of effects relate to EU firms, while a much larger share of findings (33.8%) relate to US firms (including Canada). In addition, firms in developing countries¹⁶ account for 8.3% of the effects. It is likely that CSR in developing economies with less strict regulations and less efficient market mechanisms has a lower visibility implying smaller relevance for firm performance (Wang et al., 2016). Moreover, we coded effects that relate to samples of Chinese firms (13.6%) and other BRICS countries (3.2%).¹⁷ The latter category refers to samples that either include firms from all BRICS countries jointly or individual BRICS countries except China. Finally, the variable *other_countries* (26.4%) relates to findings obtained from sample firms from a diverse set of other countries, such as Singapore that cannot be included as a separate dummy variable since they only appear in a few studies (e.g., Loh et al., 2017), or effects that were generated based on multicountry samples (e.g., von Arx & Ziegler, 2014).

The third category of variables captures the influence of CSP's and CFP's measurement as well as the set of control variables (x) included in Equations (1)–(4). There is evidence that different measurement approaches are the source of some of the variations in the results on the CSP–CFP relationship (e.g., Galant & Cadez, 2017; Wang et al., 2016). Therefore, we assess the impact that the type of CSP indicator used has on $\hat{\beta}_{CSP}$. Most estimations (82.0%) are based on CSP indices from commercial databases, such as Thomson Reuter's Asset4 or the Kinder, Lydenberg Domini (KLD)-index. Furthermore, we follow van Beurden and Gössling (2008) and assess the impact of using CSP information disclosed by companies (17.4%) or collected via questionnaires (0.6%) (cf. Table A2). This allows to identify potential bias in the estimates due to self-reported CSP (Dixon-Fowler et al., 2013). Note that in some cases (5.3%), CSP was measured by a binary variable indicating, for example, whether a firm is listed in a sustainability index (e.g., Becchetti et al., 2008). As this measurement of CSP is potentially less accurate, we use a dummy variable *CSP binary* to assess its influence on the CSP–CFP correlation. Finally, we control for the effect of using a squared CSP variable (1.9%), lagging CSP by one or more periods (17.5%), and whether an interaction between CSP and a control variable is included (10.6%) in line with Equation (3).

The operationalization of financial performance in Equations (1)–(4) is also a potential driver of heterogeneity in reported CSP–CFP effects (e.g., Galant & Cadez, 2017; Wang et al., 2016). In line with previous meta-studies, we control whether an accounting- or market-based measure was used (e.g., Hang et al., 2019). 62.1% of $\hat{\beta}_{CSP}$ were estimated using accounting-based measures, such as return-on-assets or equity (e.g., Waddock & Graves, 1997). The remaining estimates were generated using market-based measures, such as Tobin's Q (Lee et al., 2013).

In more than half of the CFP–CSP studies, CFP represents the dependent and CSP the independent variable (Equations 1–3). However, in line with slack resource theory, CFP is used as the independent variable in 45.6% of cases to predict CSP (e.g., Waddock & Graves, 1997). We employ a dummy variable (*CSP dependent*) to assess the extent to which this specification leads to different $\hat{\beta}_{CSP}$ s. We also perform several robustness checks by estimating separate models for coefficients generated using Equations (1)–(3) with CFP as dependent variable and coefficients generated based on Equation (4) where CSP is the dependent variable.

Primary regressions include either one (Equation 1) or several (Equation 2) CSP variables that capture different CSR dimensions. For example, it has been shown that the social dimension is much more visible to stakeholders than environmental performance, leading to a stronger impact on CFP (Busch & Friede, 2018). The next category of variables accounts for the dimension of the CSP indicators applied. This is important since CSR is a multidimensional construct, which implies that in the case of Equation (1), the single CSP variable can include a heterogeneous set of combinations of CSR dimensions (Wang et al., 2016). Therefore, in Equation (1), CSP is either

measured by including all three CSR dimensions (31.1%), the environmental dimension only (18.3%), the social dimension only (22.0%), the corporate governance dimension only (10.0%), or as a construct encompassing two CSR dimensions (see Table A2). In contrast, primary regressions based on Equation (2) include more than one CSP variable leading to several $\hat{\beta}_{CSP}$'s each reflecting the effect of an individual CSR dimension on CFP. For example, if a specification of Equation (2) includes a CSP variable capturing the environmental dimension and a second variable to proxy the social dimension, this leads to two $\hat{\beta}_{CSP}$'s where for the first, the *only_env* dummy, has a value of one while the *only_soci* dummy is equal to one for the second $\hat{\beta}_{CSP}$. In addition, we include a dummy variable *Csp_control* with a value of one if the primary regression is based on Equation (2), that is, controlling for additional CSP dimension(s) (37.3%) and zero otherwise (i.e., Equation 1).

The omission of variables related to CFP and CSP could lead to biased results on the CSP–CFP link (Allouche & Laroche, 2005). These variables are reflected by the vector x in Equations (1)–(4) and mainly stand for factors that relate empirically and theoretically to CFP and CSP (e.g., Goddard et al., 2005; Hirsch & Gschwandtner, 2013). A general pattern can be observed in most studies controlling for firm- and industry-specific characteristics (Lu et al., 2014) such as firm size (94.0%) (Orlitzky, 2001; Waddock & Graves, 1997) and industry-specific effects (59.9%) (Baird et al., 2012; Feng et al., 2017). 66.3% of regressions also include firms' financial risk (Feng et al., 2017). Further control variables frequently used involve the level of firms' R&D expenditures (e.g., Hull & Rothenberg, 2008) (34.4%), debt (88.1%), advertising expenditure (11.3%) (e.g., Callan & Thomas, 2009), firm age (24.9%) firms' shareholder and board characteristics (40.1 and 34.0%), and additional control variables for firm profitability (39.4%).

The final category of indicators accounts for the sample structure and estimation technique applied in the primary regressions (Hirsch, 2018; Wang & Shailer, 2015). The inclusion of the variable *number_of_years* permits us to assess whether the CSP–CFP effect is moderated by the time length of the sample analyzed. Table A2 shows that $\hat{\beta}_{CSP}$ are generated based on samples with an average time-series dimension of 9.2 years. We also include dummy variables to distinguish between effects from cross-sectional (7.6%) (e.g., Waddock & Graves, 1997) and panel studies (92.4%) (e.g., Aouadi & Marsat, 2016; Fischer & Sawczyn, 2013; Lee et al., 2013) since panel data yield more accurate estimates and realistic models (Verbeek, 2012). In addition, the econometric approach applied to estimate the CSP–CFP link potentially influences its magnitude. Thus, misspecification errors¹⁸ can be detected by including dummy variables for the estimator applied. Table A2 reveals that 49.9% of $\hat{\beta}_{CSP}$ are generated by OLS while panel estimators, such as fixed effects or GMM, are applied in 22.1% and 10.3% of the cases, respectively. Moreover, 6.7% of effects are generated with a two-stage least squares (2SLS) estimator to account for endogeneity (e.g., Fatemi et al., 2018) allowing us to identify potential bias caused by neglecting endogeneity. Finally, other methods, such as maximum likelihood estimation, were used in 11.0% of cases.

5 | RESULTS

The results of the full MRA models are reported in Table 2. As base category, we use effects from non-peer-reviewed articles that are based on Equation (1), estimated for US firms from multiple industries using a commercial CSP index that includes all CSP dimensions, a market-based CFP measure, the OLS estimator, and cross-sectional data. The results for the WLS estimation in column (1) confirm the presence of publication bias. The publication bias-corrected “true” $\hat{\beta}_{CSP}$ for the base category, that is, when all research design characteristics are set equal to zero, is 0.006 and hence in line with previous MRAs that generally detect small positive CSP–CFP effects (e.g.,

TABLE 2 MRA results with full set of research design characteristics

Variable label	WLS (SE)		BMA		BMA		BMA		Post-BMA WLS		FMA	
	Coef.	SE	Post mean	Post SD	PIP	Coef.	SE	Coef.	SE	Coef.	SE	
Constant	0.006	0.033				0.003	0.015	0.000	1.915			
se	0.408**	0.183	0.860	1.813	1.000	0.436**	0.169	0.467**	0.214			
Journal	-0.021	0.030	-5.37e-5	0.001	0.013			0.000	0.017			
Impact factor	0.004	0.008	0.001	0.002	0.184			0.000	0.004			
Last year	0.001	0.001	4.98e-7	2.07e-5	0.010			0.000	6.00e-04			
Industrials	0.019*	0.011	0.003	0.007	0.217			0.005	0.016			
Financials	0.015	0.015	3.03e-4	0.002	0.028			0.000	0.015			
Consumer goods	0.048	0.029	0.044	0.008	1.000	0.042	0.030	0.014	0.037			
EU	0.002	0.009	1.48e-4	0.001	0.025			0.000	0.002			
Developing	0.011	0.017	3.05e-4	0.002	0.026			0.000	0.012			
China	0.041**	0.016	0.044	0.002	1.000	0.044***	0.015	0.042***	0.008			
BRICS	-0.015	0.015	-4.69e-4	0.003	0.029			0.000	0.014			
Other_countries	-0.004	0.006	-1.76e-4	0.001	0.049			0.000	0.005			
CSP disclosed	-0.018	0.015	-0.007	0.008	0.502	-0.015	0.012	0.000	0.019			
CSP questionnaire	0.046**	0.023	0.007	0.017	0.156			0.015	0.037			
CSP binary	-0.033***	0.010	-0.031	0.004	1.000	-0.030***	0.009	-0.029***	0.006			
CSP_squared	-0.047***	0.009	-0.044	0.008	1.000	-0.044***	0.008	-0.033***	0.020			
CSP_lag	-0.011**	0.005	-0.012	0.002	1.000	-0.012***	0.004	-0.004	0.009			
CSP_interaction	-2.76e-4	0.005	4.261e-6	2.15e-4	0.007			0.000	3.00e-04			
CFP accounting	0.005	0.005	0.005	0.002	0.882	0.006	0.005	0.000	0.005			
CSP dependent	-0.017***	0.006	-0.018	0.002	1.000	-0.019***	0.006	-0.010*	0.008			
Only_gov_CSP	-0.003	0.019	-4.50e-6	3.04e-4	0.008			0.000	0.003			
Only_env_CSP	0.004	0.009	1.31e-5	2.40e-4	0.010			0.000	0.004			

(Continues)

TABLE 2 (Continued)

Variable label	WLS (SE) (1)		BMA (2)		Post SD	PIP	Post-BMA WLS (3)		FMA (4)	
	Coef.	SE	Post mean	SE			Coef.	SE	Coef.	SE
Only_soc_CSP	0.004	0.006	-4.51e-4		1.95e-4	0.010		0.000	0.004	
Soc&Env_CSP	0.002	0.005	8.48e-6		1.82e-4	0.010		0.000	0.002	
Env&Gov_CSP	0.012	0.012	9.13e-5		0.002	0.009		0.000	0.013	
Soc&Gov_CSP	0.004	0.006	2.99e-5		0.001	0.009		0.000	0.004	
CSP_control	-0.008*	0.005	-0.007		0.002	0.973	-0.008**	0.004	0.008	
Size_control	0.032***	0.016	0.034		0.003	1.000	0.032**	0.015	0.030***	
Industry_control	0.004	0.004	0.001		0.002	0.193		0.000	0.005	
Risk_control	-0.008	0.006	-0.001		0.002	0.250		0.000	0.008	
R&D_control	0.003	0.005	0.001		0.002	0.252		0.000	0.003	
Debt_control	0.007	0.016	-2.09e-6		3.34e-4	0.010		0.000	0.008	
Advertising_control	0.003	0.005	6.30e-5		0.001	0.023		0.000	0.003	
Age_control	0.004	0.006	0.001		0.002	0.245		0.000	0.004	
Profit_control	-0.010*	0.006	-0.009		0.002	1.000	-0.008	0.005	-0.007**	
Share_control	0.007	0.005	0.008		0.002	1.000	0.008*	0.005	0.006	

(Continues)

TABLE 2 (Continued)

Variable label	WLS (SE)		BMA		BMA		BMA		Post-BMA WLS		FMA	
	Coef.	SE	SE	Post mean	Post SD	PIP	Coef.	SE	Coef.	SE	Coef.	SE
Board_control	0.002	0.006	0.006	8.15e-5	0.001	0.031					0.000	0.002
Number of years	-4.45e-4	0.001	0.001	-6.04e-6	4.36e-5	0.027					0.000	5.00e-04
Cross	-0.011	0.011	0.011	-6.96e-5	0.001	0.011					0.000	0.011
Fixed	-0.020***	0.005	0.005	-0.021	0.002	1.000	-0.021***	0.005	-0.021***	0.005	-0.021***	0.005
GMM	-0.022***	0.007	0.007	-0.020	0.003	1.000	-0.021***	0.005	-0.017**	0.005	-0.017**	0.008
2SLS	-0.014**	0.006	0.006	-0.013	0.003	0.999	-0.013**	0.005	-0.011***	0.005	-0.011***	0.004
Other_estimator	-0.009	0.008	0.008	-0.005	0.005	0.478					0.000	0.009
Obs.	7800			7800			7800				7800	
R ²	0.134						0.124					
F-value	6.35***						6.82***					
Studies	512			512			512				512	

Note: Study-cluster robust standard errors in parentheses. Base group: *Non peer-reviewed, mixed industries, US, CSP-Index, CFP-market, All_CSP, cross, OLS*. Post-BMA regressions are performed using variables with PIP > 0.5.

Abbreviation: PIP, "posterior inclusion probability."

*** $p < .01$, ** $p < .05$, * $p < .1$.

Hang et al., 2019). Since the constant only reflects the “true” effect beyond publication bias, if all study design characteristics are set equal to zero (i.e., for the base group), it should be interpreted in the context of the significant covariates related to the study design. We discuss these results based on some best practice examples below.

Several research design characteristics are found to affect heterogeneity in $\hat{\beta}_{CSP}$. In contrast to Busch and Friede's (2018) MRA, no differences can be detected between non-peer-reviewed results and those published in journals. This confirms our assumption that working papers are written with the intention of publishing them. Unlike Busch and Friede (2018), our MRA does not reveal different effects in journals with an IF compared with journals without an IF such that journal quality does not seem to moderate $\hat{\beta}_{CSP}$. Several meta-studies on the CSP–CFP effect found that they decrease over time, which can be explained by learning effects (Allouche & Laroche, 2005; Lu & Taylor, 2016). However, in line with the meta-analysis by Friede et al. (2015), our findings reveal stable effects over time. As expected, the relationship between CSP and CFP is more pronounced in studies that focus exclusively on industrial sectors (+0.019) as these are potentially associated with stronger environmental effects. These findings also suggest that a multi-industry focus may obscure industry-specific effects leading to a lower CSP–CFP relationship (Allouche & Laroche, 2005). Moreover, we find that regressions focussing specifically on firms operating in China generate $\hat{\beta}_{CSP}$ s that are 0.041 higher compared to the base category of US firms. This indicates that the CSP–CFP relationship is moderated by differences in firms' CSR practices across countries (e.g., Matten & Moon, 2008; Wang et al., 2016). These findings are also in line with previous meta-analyses, which generally detect smaller effects for US samples (e.g., Allouche & Laroche, 2005; Lu & Taylor, 2016). In contrast to Wang et al. (2016), we do not identify lower effects for developing countries in comparison to the United States.

As regards the measurement of CFP, some meta-studies show that accounting-based measures are more strongly related to CSP, since they best reflect the internal capabilities of firms (e.g., Busch & Friede, 2018; Huang, 2019). However, differences to market-based measures are generally small (e.g., Vishwanathan et al., 2020) with some meta-studies even detecting a stronger link of CSP to market-based CFP (e.g., Dixon-Fowler et al., 2013). Accordingly, we do not find a difference between using an accounting-based measure and market-based measures.

In line with previous meta-studies, CSP measures derived from questionnaires entail higher $\hat{\beta}_{CSP}$ (+0.046) compared to cases where a commercial CSP index is used (Hou et al., 2016; Wang et al., 2016). Oppositely, our results show that binary CSP measures lead to $\hat{\beta}_{CSP}$ s being 0.033 lower compared to applications of a CSP index measure. As binary CSP measures only indicate whether firms perform CSR in general, $\hat{\beta}_{CSP}$ estimates might be less accurate. Moreover, using a squared or lagged CSP measure leads to lower effects by 0.047 and 0.011, respectively. Regarding the operationalization of CSP dimensions, we do not detect any difference between analyses that employ multidimensional (including several of the three CSR dimensions) or one-dimensional CSP variables. However, controlling for other CSP dimensions leads to lower effects (−0.008). We also find that the type of theoretical framework tested can affect the resulting $\hat{\beta}_{CSP}$ since models relying on the slack resource theory (Equation 4), that is, with CSP as dependent variable generate lower effects (−0.017) compared to studies where CFP is the dependent variable (Equations 1–3).

Allouche and Laroche (2015) find in their MRA that including control variables such as size, industry, risk, or R&D does not alter estimated CSP–CFP effects. In contrast, we find that controlling for firm size leads to significantly higher estimates (+0.032) while including an additional proxy for firm profitability leads to significantly lower effects (−0.010). Finally, Lu and Taylor (2016) report that more complex econometric approaches, such as 2SLS, yield smaller $\hat{\beta}_{CSP}$ in line

with Busch and Friede (2018). We also find that econometric specifications such as fixed effects, GMM, and 2SLS generate lower $\hat{\beta}_{CSP}$ effects compared to OLS.

We apply several robustness checks for the full MRA model. Table A5 shows that the results are only marginally affected by the approach to assess publication bias such as splitting the standard error between positive and negative effects or using the number of observations as an instrument for the standard error. Moreover, the robustness checks show that using wild bootstrapped standard errors leads to somewhat lower significances and that our results are not affected by the exclusion of observations for which only asterisks are available as precision measures. However, we detect some differences in the drivers of heterogeneity in $\hat{\beta}_{CSP}$ when only observations that relate to one of the specifications in Equations (1)–(3), that is, with CFP as the dependent variable, are considered (Table A5, column 3). For these observations, we detect a higher relationship between CSP and CFP beyond publication bias for the base group (0.084). Moreover, differences by the choice of industry analyzed, including firm size as a control variable and lagging the CSP variable does not affect $\hat{\beta}_{CSP}$ estimates in this case. However, it can be observed, that samples with a longer time series dimension generate on average lower $\hat{\beta}_{CSP}$ effects (−0.002 per additional year) in this case while using samples for firms in developing countries leads to lower effects (−0.028). The latter finding is in line with Wang et al. (2016) and potentially caused by the fact that CSR in developing economies with less strict regulations and less efficient market mechanisms has a lower visibility and thus relevance for firm performance.

Regarding model diagnosis, the F -test reveals the overall significance of all models and the R^2 indicates that up to 32% of the variation in $\hat{\beta}_{CSP}$ can be ascribed to publication bias and research design characteristics in the model that accounts for bi-directional publication bias (Table A5, column 1).

The fact that there is no theoretical foundation for the inclusion of individual research design variables (k_j) in Equation (12) raises issues of model uncertainty (e.g., Hang et al., 2019; Havranek et al., 2017). Therefore, we use BMA and FMA techniques as additional robustness checks to identify the independent variables with explanatory power for the variance in reported CSP–CFP correlations (De Luca & Magnus, 2011; Raftery et al., 1997; Zigraviova et al., 2021). BMA computes the weighted average of the conditional estimates from a set of models including all possible combinations of potentially relevant regression coefficients. We set the standard error as a fixed regressor, which leaves us with 41 potentially relevant covariates, from which approximately 4.4 trillion possible models can be constructed. However, computational constraints obliged us to limit the BMA to 1 mio. Variable draws leading to 125,585 inspected models with an average of 17.68 variables included (see the online appendix for more information on the BMA). BMA weights each of the resulting models and the estimates of its parameters using data and priors to provide averages of post mean values and standard errors. Note that we use unit information g-prior where the prior receives the same importance as an individual observation and random model prior for model probability (e.g., Havranek et al., 2018; Zeugner & Feldkircher, 2015). The posterior inclusion probability (PIP) that a variable is included in the model can be used as an indicator for its relevance (De Luca & Magnus, 2011; Fernau & Hirsch, 2019). Following Zigraviova and Havránek (2016), we only include covariates with a PIP larger than 0.5 in a reduced post-BMA OLS model.¹⁹ The BMA results are graphically summarized in Figure 4, which displays the importance of individual study-design characteristics for the heterogeneity in reported CSP–CFP coefficients. On the y -axis, the variables are ordered by their PIP while the x -axis displays individual models ordered by posterior model probabilities. Hence the best models regarding data fit are found on the left. Blue and red colors indicate that a variable has a positive or negative effect, respectively (e.g., Zigraviova et al., 2021). Numerical BMA results are shown in Table 2 (column 2) and reveal that,

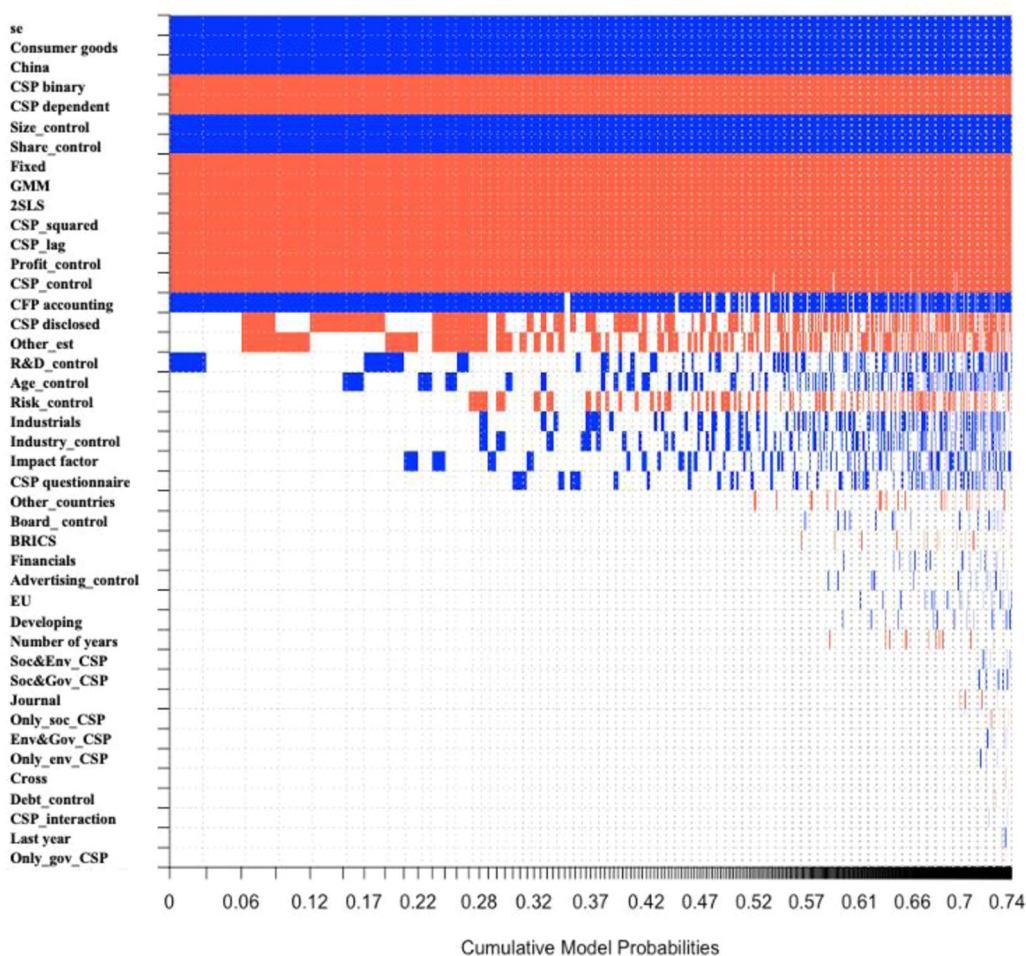


FIGURE 4 BMA model inclusion based on best 500 models. *Note:* The figure displays the importance of individual study-design characteristics for the heterogeneity in reported CSP–CFP coefficients based on BMA. On the y-axis the variables are ordered by their PIP while the x-axis displays individual models ordered by posterior model probabilities. Blue and red colors indicate that a variable has a positive or negative effect, respectively. [Colour figure can be viewed at wileyonlinelibrary.com]

besides the standard error, 15 research design covariates arrive at PIPs exceeding 0.5, which are, therefore, included in the post-BMA MRA model. The post-BMA regression in column 3 confirms the results of the full WLS model (column 1). Since BMA results may vary depending on the prior information, we also apply FMA, which does not rely on priors. We follow the R procedure proposed in Havranek et al. (2017) to apply the FMA. FMA is based on Mallows' criteria to generate asymptotically optimal weights (Amini & Parmeter, 2012; Hansen, 2007; Zigraviova et al., 2021). The results are reported in column 4 of Table 2 and mainly resemble the post-BMA. Finally, we applied BMA and FMA using the number of observations as an instrument for the standard error without finding changes in the results (c.f. Table A5 columns 5–7).

Thus far, we have observed that the average reported CSP–CFP effect beyond publication bias is close to zero for the base group (effects from non-peer-reviewed articles that are based on Equa-

TABLE 3 Best practice results

	EU	US	China	BRICS	Developing
Best-practice baseline	0.034* (0.020)	0.032* (0.017)	0.073*** (0.021)	0.017 (0.022)	0.043* (0.023)
Manufacturing industry	0.053** (0.021)	0.051*** (0.018)	0.092*** (0.022)	0.036 (0.023)	0.062** (0.025)
CSP dependent	0.036 (0.022)	0.034* (0.020)	0.075*** (0.023)	0.019 (0.024)	0.045* (0.026)
Fixed-effects	0.014 (0.022)	0.012 (0.019)	0.053** (0.022)	-0.003 (0.023)	0.023 (0.024)
Last year of analysis 1990	0.035* (0.020)	0.033** (0.016)	0.074*** (0.021)	0.018 (0.021)	0.043* (0.022)
Last year of analysis 2020	0.051*** (0.016)	0.049*** (0.016)	0.090*** (0.022)	0.034 (0.021)	0.060** (0.024)
Last year of analysis 2020 & fixed effects	0.031** (0.016)	0.029* (0.015)	0.070*** (0.021)	0.014 (0.019)	0.040* (0.022)

Note: Standard errors in parentheses.

*** $p < .01$, ** $p < .05$, * $p < .1$.

tion 1, estimated for US firms from multiple industries using a commercial CSP index that includes all CSP dimensions, a market-based CFP measure, the OLS estimator, and cross-sectional data). The average effect for the base group can serve as a basis to further evaluate the influence of the study design characteristics revealed by the MRA. We construct several “best-practice” scenarios for the study design and estimate the resulting effects across industries and regions. We assume that “best practice” in empirical CSP–CFP research involves publication in a peer-reviewed journal with an IF, the use of panel data and a continuous CSP indicator from a commercial index, as well as the inclusion of size, industry dummies, and debt as most frequently used control variables. The best-practice results are presented in Table 3 and yield small positive partial correlations between 0.017 and 0.034 across the five countries. The results change only marginally when the definition of best practice is changed to focusing on a single industry such as manufacturing only. Nevertheless, assuming that best practice involves using a panel estimator such as fixed effects the partial correlation becomes insignificant for all countries except China. Moreover, we find that estimating models with CSP as dependent variable lowers the effect. Finally, if we assume that best practice also involves using recent data samples, we observe somewhat higher effects. To summarize, it must be acknowledged that predicted $\hat{\beta}_{CSP}$ are generally small values with low economic importance (Cohen, 1988) and are thus apt to be irrelevant for formulating suggestions on firms’ CSR strategies.

6 | SUMMARY AND CONCLUSIONS

Our MRA of 512 empirical studies with 7800 correlations ($\hat{\beta}_{CSP}$) on the relationship between CSP and CFP published in the last three decades points to a small positive link between both measures. We take the average of the 10% most precisely estimated correlations to obtain a proxy for the “true” CFP–CSP effect beyond publication bias of 0.023. This value is lower than the simple sample mean of 0.035 pointing to a moderate degree of publication bias. A descriptive and graphical analysis of the variation in $\hat{\beta}_{CSP}$ values reveals excess variation around the proxy of the “true” correlation. This is an indicator of bi-directional publication bias that favors either statistically significant positive or negative $\hat{\beta}_{CSP}$.

The degree of publication bias and the presence of a genuine effect is assessed more formally by conducting a MRA using WLS regression. In accordance with the descriptive analysis, the MRA results confirm that the variation in reported estimates is influenced by publication bias. This

suggests a tendency to select statistically significant results in publications. Moreover, the MRA reveals that after correcting for publication bias, a small positive CSP–CFP effect of 0.017 prevails. We also apply recently developed nonlinear approaches to detect publication bias. Those yield values between 0.017 and 0.045 depending on the method applied with an average mean effect of 0.027. Some of those results are closer to the sample mean of 0.035, which would suggest that the degree of publication bias is marginal.

Adding variables related to the underlying research designs to the MRA reveals the significance of several characteristics that affect the CSP–CFP relationship. We find that a significant increase in $\hat{\beta}_{CSP}$, by a value of 0.041, emerges when the focus is directed at samples of firms from China instead of US firms. Furthermore, when studying industrial branches exclusively, the $\hat{\beta}_{CSP}$'s exceed those reported in multi-industry studies by 0.019 whereas we do not find any significant difference between financial and consumer sector compared to multi-industry analyses. Our results imply that it is important that future CSP–CFP studies only interpret results and derive related implications after giving due consideration to the underlying study focus. This is essential when comparing results of studies with a different geographical, or industry focus. Moreover, a correct specification of the underlying regression model used to estimate the CSP–CFP relationship is also important. In this context, the MRA shows that the use of a binary CSP proxy, which is likely to be less accurate, or omitting firms' size and financial performance as control variables significantly affects the resulting $\hat{\beta}_{CSP}$. It follows that future research should avoid less accurate $\hat{\beta}_{CSP}$ estimates by using continuous indicators such as KLD or Asset4 and implement a comprehensive selection process based on relevant theoretical approaches for the inclusion of control variables that are related to CSP and CFP. Moreover, the way in which the CSP variable is defined can also alter the generated findings. In that respect, our MRA reveals, that using a squared or lagged CSP variable as well as including additional CSP dimensions as control variables leads to significantly smaller $\hat{\beta}_{CSP}$ effects. Hence, researchers should carefully interpret their results accounting for the estimated functional relationship between CSP and CFP and the CSR dimension(s) considered.

Regarding technical study characteristics, we find that estimating the relationship with the “right” econometric approach is an important issue. While we cannot draw any conclusions about the adequate econometric estimator to use, the results show that the application of OLS ignoring the panel structure of the data or potential endogeneity in the CSP–CFP relationship can lead to upward biased results compared to fixed effects, GMM, and 2SLS.

Finally, our MRA provides evidence that it is important to interpret results in line with the theoretical approach tested. The MRA demonstrates that studies, which apply a framework with CSP as dependent variable (Equation 4) to test for slack resource theory generate significantly smaller $\hat{\beta}_{CSP}$ effects compared to analyses where CFP is used as a dependent variable (Equations 1–3). Moreover, we find that the heterogeneity of effects generated by both research streams is driven by a different set of study design variables.

The main limitation of the MRA framework is that it is not possible to detect misspecification, that is common to all studies. For example, there might be selection bias towards firms that engage in CSP (Shahzad & Sharfman, 2017) or based on the size of the firms analyzed (Hou et al., 2016). While Orlitzky (2001) shows that firm size does not confound the CSP–CFP relationship, most commercial databases on firms' financial information lack details about very small firms for which the effect of CSR on CFP may be smaller. This can subsequently lead to an overestimation of the “true” CSP–CFP correlation. Finally, no matter how carefully all potentially important moderators of the CSP–CFP effect and the application of BMA are considered, the MRA might fail to capture additional, unobservable drivers of the relationship, such as researchers' academic background or working style. Moreover, there is a trade-off between controlling for the largest number

of study design characteristics possible and not coding dummy variables that are only related to a small fraction of the estimates, which could cause econometric issues in the estimation of the MRA. This might also be the reason why the MRA can only explain up to 32% of the variance in reported $\hat{\beta}_{CSPs}$, a value which is in line with previous MRAs on the CSP–CFP effect (Allouche & Laroche, 2005; Huang et al., 2020). We nonetheless believe that our MRA represents a meaningful starting point to achieve a higher degree of transparency across an important research stream. It can provide guidelines for the correct implementation of underlying regressions and highlights the importance of giving due consideration to the study setting when interpreting results. Finally, our results may help to prevent future misuse of statistical practices, particularly in the context of the ongoing p -value debate (Goodman, 2019; Greenland, 2019).

ACKNOWLEDGMENTS

The author would like to thank three anonymous referees as well as the associate editor for extremely valuable comments on the article.

Open Access funding enabled and organized by Projekt DEAL.

WOA Institution: UNIVERSITÄT HOHENHEIM


Consortia Name: Projekt DEAL

ORCID

Stefan Hirsch  <https://orcid.org/0000-0003-2668-6824>

Thies Petersen  <https://orcid.org/0000-0002-1870-9603>

Maximilian Koppenberg  <https://orcid.org/0000-0002-5814-7836>

Monika Hartmann  <https://orcid.org/0000-0002-3631-164X>

ENDNOTES

¹Our MRA is based on the guidelines by Stanley et al. (2013) and Havránek et al. (2020).

²MRA was originally applied in natural sciences, for example, to summarize clinical trials (e.g., DerSimonian & Laird, 1986). Since the 1990s, starting with the seminal study on the effect of minimum wages by Card and Krueger (1995), it has been applied increasingly in economic research (e.g., Oczkowski & Doucouliagos, 2015). Recent MRAs include, for example, Bajzik et al. (2020) who analyze the substitution elasticity between imported and domestic goods, Balima et al. (2020) who investigate the relevance of publication bias for macroeconomic effects of inflation targeting, Matousek et al. (2022) who focus on individual discount rates, Gerchert et al. (2022) and Knoblach et al. (2020) who evaluate the substitution elasticity between labor and capital or Zigraiova et al. (2021) who focus on the relationship of forward and spot rates.

³According to Carroll, CSR captures “the general notion of businesses’ accountability or responsibility to society.” (Carroll, 2018, p. 746) while CSP extends the concept of CSR and focuses “on actual results achieved.” (Carroll, 2018, p. 746) with only the latter being measurable. We follow Carroll and use the terms CSR and CSF, accordingly.

⁴Note that the debate on how to counter the abuse of p -values and the nonreplicability of findings has recently intensified with the *American Statistical Association* (ASA) publishing a special issue on how to improve statistical practice and moving to a world where inference is not based solely on p -values falling below a specific threshold (Wasserstein and Lazar, 2016; Hubbard, 2019; McShane et al. 2019; Wasserstein et al. 2019).

⁵Three of the main authors and a research assistant screened and coded the literature. The final dataset was then crosschecked for inconsistencies and, if necessary, corrected after consultation among the main authors until full agreement on the coding was achieved.

⁶Note that we did not limit our literature search to a specific timeframe and the 512 articles were identified solely on the basis of the inclusion criteria described above.

⁷In the case of studies, which did not report t -statistics or standard errors of regression coefficients but only p -values, we converted p -values to t -values considering the regression’s dfs (Wang & Shailer, 2015).

- ⁸Note that the distribution of partial correlation coefficients is usually not normal close to the boundary values of -1 and 1 . Fisher's transformation can be applied (e.g., Wang & Shailer, 2015; Oczkowski & Doucouliagos, 2015) to correct this. However, Stanley and Doucouliagos (2012) conclude that this transformation has little practical effect on the MRA results, and therefore, we refrain from applying it in the present case.
- ⁹The null hypothesis of no skewness is rejected at the 5%-level. Histograms for β_{CSP} related to individual CSP dimensions can be found in the appendix Figure A1.
- ¹⁰Figure A1 presents a histogram if cases where only asterisks are reported are excluded and A2 displays a funnel plot where such observations are highlighted in different color.
- ¹¹Similar results can be observed when the funnel graphs are inspected for partial correlations related to the subgroups of CSP dimensions depicted in A2.
- ¹²In our sample, the size of study-clusters ranges from 1 to 296 observations.
- ¹³There is a trade-off between controlling for all relevant study design characteristics and not having dummy variables that are only related to a small fraction of the estimates. While the former is essential to reduce biases due to unobserved methodological issues dummy variables that only relate to a very small fraction of the estimates would involve econometric issues. We do our best in finding the right mixture between both aspects. Moreover by controlling for the publication year, we capture the effect of methodological innovations in newer studies that are not coded as individual dummy variables (Havranek et al., 2018).
- ¹⁴It would be interesting to compare the working paper results that are omitted with those that actually appear later in the published version. However, as indicated above, we only identified one published article with omitted working paper results.
- ¹⁵According to Matten and Moon (2008), "explicit" CSR clearly states responsibility for certain societal interests, is based on corporate discretion, and thus reflects neither governmental authority nor broader formal or informal institutions. "Implicit" CSR refers to the role of corporations within society and reflects formal and informal societal institutions. Implicit CSR is embedded in societal values, norms, and rules, which implies that corporations have an obligation to consider stakeholders' issues.
- ¹⁶We use the World Bank definition for classifying developing countries.
- ¹⁷BRICS stands for the five emerging economies: Brazil, Russia, India, China, and South Africa. As we consider China separately, BRICS refers in our study to the four remaining countries unless a joint sample of firms from BRICS countries is used.
- ¹⁸For example, when pooled OLS is applied to a panel dataset instead of a more suitable panel estimator.
- ¹⁹Effects are considered as weak for PIPs between 0.5 and 0.75, positive for 0.75–0.95, strong for 0.95–0.99, and decisive for 0.99–1 (Havranek et al., 2017).

DATA AVAILABILITY STATEMENT

All meta data and estimation codes as well as supplementary material are shared via the journal web page.

REFERENCES

- van Aert, R. C., & van Assen, M. (2021). *Correcting for publication bias in a meta-analysis with the p-uniform* method* [Working paper]. Tilburg University.
- Allouche, J., & Laroche, P. (2005). A Meta-analytical investigation of the relationship between corporate social and financial performance (pp. 18). *Revue de Gestion des Ressources Humaines*, Eska. hal-00923906.
- Amoako, G. K., & Dartey-Baah, K. (2020). Corporate social responsibility: Strategy for boosting brand perception and competitive advantage. In A. Kuna-Marszalek & A. Klysik-Uryszek (Ed.), *CSR and socially responsible investing strategies in transitioning and emerging economies*. IGI Global.
- Amini, S. M., & Parmeter, C. F. (2012). Comparison of model averaging techniques: Assessing growth determinants. *Journal of Applied Econometrics*, 27(5), 870–876.
- Andrews, I., & Kasy, M. (2019). Identification of and correction for publication bias. *American Economic Review*, 109(8), 2766–2794.
- Angrist, J. D., & Pischke, J. - S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

- Aouadi, A., & Marsat, S. (2016). Do ESG controversies matter for firm value? Evidence from international data. *Journal of Business Ethics, 15*(2), 1–21.
- Baird, P. L., Geylani, P. C., & Roberts, J. A. (2012). Corporate social and financial performance re-examined: Industry effects in a linear mixed model analysis. *Journal of Business Ethics, 109*(3), 367–388.
- Bajzik, J., Havranek, T., Irsova, Z., & Schwarz, J. (2020). Estimating the Armington elasticity: The importance of study design and publication bias. *Journal of International Economics, 127*, 103383.
- Balima, H. W., Kilama, E. G., & Tapsoba, R. (2020). Inflation targeting: Genuine effects or publication selection bias? *European Economic Review, 128*, 103520.
- Barnett, M. L., & Salomon, R. M. (2012). Does it pay to be really good?: Addressing the shape of the relationship between social and financial performance. *Strategic Management Journal, 33*(11), 1304–1320.
- Becchetti, L., Di Giacomo, S., & Pinnacchio, D. (2008). Corporate social responsibility and corporate performance: Evidence from a panel of US listed companies. *Applied Economics, 40*(5), 541–567.
- Belu, C., & Manescu, C. (2013). Strategic corporate social responsibility and economic performance. *Applied Economics, 45*(19), 2751–2764.
- Billor, N., Hadi, A. S., & Velleman, P. F. (2000). BACON: Blocked adaptive computationally efficient outlier nominators. *Computational Statistics & Data Analysis, 34*, 279–298.
- Bom, P. R. D., & Ligthart, J. E. (2009). *How productive is public capital? A meta-regression analysis* [Working paper 09-12]. International Studies Program, Georgia State University.
- Bom, P. R. D., & Rachinger, H. (2019). A kinked meta-regression model for publication bias correction. *Research Synthesis Methods, 10*(4), 497–514.
- Brower, J., & Dacin, P. A. (2020). An institutional theory approach to the evolution of the corporate social performance–corporate financial performance relationship. *Journal of Management Studies, 57*(4), 805–836.
- Busch, T., & Friede, G. (2018). The robustness of the corporate social and financial performance relation: A second-order meta-analysis. *Corporate Social Responsibility and Environmental Management, 25*, 583–608.
- Callan, S. J., & Thomas, J. M. (2009). Corporate financial performance and corporate social performance: An update and reinvestigation. *Corporate Social Responsibility and Environmental Management, 16*(2), 61–78.
- Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics, 90*(3), 414–427.
- Capaldi, N., Idowu, S. O., & Schmidpeter, R. (Eds.). (2017). *Dimensional corporate governance*. Springer International Publishing.
- Capelle-Blancard, G., & Petit, A. (2017). The Weighting of CSR Dimensions: One Size Does Not Fit All. *Business and Society, 56*(6), 919–943.
- Card, D., & Krueger, A. B. (1995). Time-series minimum-wage studies: A meta-analysis. *American Economic Review, 85*, 238–243.
- Carroll, A. B. (2018). Corporate social responsibility (CSR) and corporate social performance (CSP). In R. W. Kolb (Ed.), *The SAGE encyclopedia of business ethics and society* (pp. 746–754). Sage Publications.
- Cavaco, S., & Crifo, P. (2014). CSR and financial performance: Complementarity between environmental, social and business behaviours. *Applied Economics, 46*(27), 3323–3338.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Erlbaum.
- Crifo, P., & Forget, V. D. (2015). The economics of corporate social responsibility: A firm-level perspective survey. *Journal of Economic Surveys, 29*(1), 112–130.
- Dahlsrud, A. (2008). How corporate social responsibility is defined: An analysis of 37 definitions. *Corporate Social Responsibility and Environmental Management, 15*(1), 1–13.
- Davis, K. (1973). The case for and against business assumption of social responsibilities. *Academy of Management Journal, 16*, 312–322.
- De Luca, G., & Magnus, J. R. (2011). Bayesian model averaging and weighted-average least squares: Equivariance, stability, and numerical issues. *The Stata Journal, 11*(4), 518–544.
- Dersimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials, 7*(3), 177–188.
- Dixon-Fowler, H. R., Slater, D. J., Johnson, J. L., Ellstrand, A. E., & Romi, A. M. (2013). Beyond “does it pay to be green?” A meta-analysis of moderators of the CEP–CFP relationship. *Journal of Business Ethics, 112*, 353–366.
- Doucouliaos, H., Laroche, P., & Stanley, T. D. (2005). Publication bias in union-productivity research? *Relations Industrielles/Industrial Relations, 60*(2), 320–347.

- Doucouliafos, H., & Stanley, T. D. (2009). Publication selection bias in minimum-wage research? A meta-regression analysis. *British Journal of Industrial Relations*, 47(2), 406–428.
- Dupire, M., & M'Zali, B. (2018). CSR strategies in response to competitive pressures. *Journal of Business Ethics*, 148, 603–623.
- Endrikat, J. (2016). Market reactions to corporate environmental performance related events: A meta-analytic consolidation of the empirical evidence. *Journal of Business Ethics*, 138, 535–548.
- Commission, E. (2003). *Responsible entrepreneurship: A collection of good practice cases among small and medium-sized enterprises across Europe*. European Commission.
- Commission, E. (2011). *Communication from the commission to the European Parliament, the council, the European Economic and Social Committee and the committee of the regions /A renewed EU strategy 2011–14 for Corporate Social Responsibility*. European Commission. COM(2011) 681 final.
- Fatemi, A., Glaum, M., & Kaiser, S. (2018). ESG performance and firm value: The moderating role of disclosure. *Global Finance Journal*, 38, 45–64.
- Feng, M., Wang, X., & Kreuze, J. G. (2017). Corporate social responsibility and firm financial performance. *American Journal of Business*, 32(3/4), 106–133.
- Fernau, E., & Hirsch, S. (2019). What drives dividend smoothing?: A meta regression analysis of the Lintner model. *International Review of Financial Analysis*, 61, 255–273.
- Fischer, T. M., & Sawczyn, A. A. (2013). The relationship between corporate social performance and corporate financial performance and the role of innovation: Evidence from German listed firms. *Journal of Management Control*, 24(1), 27–52.
- Freeman, E. (1984). *Strategic management: A stakeholder approach*. Pitman.
- Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4), 210–233.
- Friedman, M. (1970). The social responsibility of business is to increase its profits. *The New York Times Magazine*.
- Furukawa, C. (2019). *Publication bias under aggregation frictions: Theory, evidence, and a new correction method*. Available at SSRN: <https://ssrn.com/abstract=3362053>
- Galant, A., & Cadez, S. (2017). Corporate social responsibility and financial performance relationship: A review of measurement approaches. *Economic Research-Ekonomska Istrazivanja*, 30(1), 676–693.
- Gerchert, S., Havranek, T., Irsova, Z., & Kolcunova, D. (2022). Measuring capital-labor substitution: The importance of method choices and publication bias. *Review of Economic Dynamics*, (in press).
- Goddard, J., Tavakoli, M., & Wilson, J. O. S. (2005). Determinants of profitability in European manufacturing and services: Evidence from a dynamic panel model. *Applied Financial Economics*, 15(18), 1269–1282.
- Goodman, S. N. (2019). Why is getting rid of P-values so hard? Musings on science and statistics. *The American Statistician*, 73(s1), 26–30.
- Greene, W. H. (2012). *Econometric analysis* (7th ed.). Pearson Education.
- Greenland, S. (2019). Valid P-values behave exactly as they should: Some misleading criticisms of P-values and their resolution with S-values. *The American Statistician*, 73(s1), 106–114.
- Grewatsch, S., & Kleindienst, I. (2017). When does it pay to be good? Moderators and mediators in the corporate sustainability–corporate financial performance relationship: A critical review. *Journal of Business Ethics*, 145, 383–416.
- Hang, M., Geyer-Klingenberg, J., & Rathgeber, A. W. (2019). It is merely a matter of time: A meta-analysis of the causality between environmental performance and financial performance. *Business Strategy and the Environment*, 28(2), 257–273.
- Hansen, B. E. (2007). Least squares model averaging. *Econometrica*, 75(4), 1175–1189.
- Hartmann, M. (2011). Corporate social responsibility in the food sector. *European Review of Agricultural Economics*, 38(3), 297–324.
- Havráněk, T., Stanley, T. D., Doucouliagos, H., Bom, P., Geyer-Klingenberg, J., Iwasaki, I., Reed, W. R., Rost, K., & Aert, R. C. M. (2020). Reporting guidelines for meta-analysis in economics. *Journal of Economic Surveys*, 34(3), 469–475.
- Havranek, T., Irsova, Z., & Zeynalova, O. (2018). Tuition fees and university enrolment: A meta-regression analysis. *Oxford Bulletin of Economics and Statistics*, 80(6), 1145–1184.
- Havranek, T., Rusnak, M., & Sokolova, A. (2017). Habit formation in consumption: A meta-analysis. *European Economic Review*, 95, 142–167.

- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academia Press.
- Hirsch, S. (2018). Successful in the long run: A meta-regression analysis of persistent firms profits. *Journal of Economic Surveys*, 32(1), 23–49.
- Hirsch, S., & Gschwandtner, A. (2013). Profit persistence in the food industry: Evidence from five European countries. *European Review of Agricultural Economics*, 40(5), 741–759.
- Ho, S. S. M., Yuansha Li, A., Tam, K., & Tong, J. Y. (2016). Ethical image, corporate social responsibility, and R&D valuation. *Pacific-Basin Finance Journal*, 40(part B), 335–348.
- Hou, M., Liu, H., Fan, P., & Wei, Z. (2016). Does CSR practice pay off in East Asian firms? A meta-analytic investigation. *Asia Pacific Journal of Management*, 33(1), 195–228.
- Huang, D. Z. X. (2019). Environmental, social and governance (ESG) activity and firm performance: A review and consolidation. *Accounting and Finance*, <https://doi.org/10.1111/acfi.12569>
- Huang, K., Sim, N., & Zhao, H. (2020). Corporate social responsibility, corporate financial performance and the confounding effects of economic fluctuations: A meta-analysis. *International Review of Financial Analysis*, 70, 101504.
- Hubbard, R. (2019). Will the ASA's efforts to improve statistical practice be successful? Some evidence to the contrary. *The American Statistician*, 73(s1), 31–35.
- Hull, C. E., & Rothenberg, S. (2008). Firm performance: The interactions of corporate social performance with innovation and industry differentiation. *Strategic Management Journal*, 29(7), 781–789.
- Ioannidis, J. P. A., Stanley, T. D., & Doucouliagos, H. (2017). The power of bias in economics research. *The Economic Journal*, 127(605), 236–265.
- Kim, J., Doucouliagos, H., & Stanley, T. D. (2014). *Market efficiency in Asian and Australasian stock markets: A fresh look at the evidence* [Economics series working paper SWP 2014/9]. School of Accounting, Economics and Finance, DEAKIN University.
- Knoblauch, M., Roessler, M., & Zwerschke, P. (2020). The elasticity of substitution between capital and labour in the US economy: A meta-regression analysis. *Oxford Bulletin of Economics and Statistics*, 82(1), 62–82.
- Latapí Agudelo, M. A., Jóhannsdóttir, L., & Davídsdóttir, B. (2019). A literature review of the history and evolution of corporate social responsibility. *International Journal of Corporate Social Responsibility*, 4(1), 1–23.
- Lee, S., Seo, K., & Sharma, A. (2013). Corporate social responsibility and firm performance in the airline industry: The moderating role of oil prices. *Tourism Management*, 38, 20–30.
- Loh, L., Thomas, T., & Wang, Y. (2017). Sustainability reporting and firm value: Evidence from Singapore-listed companies. *Sustainability*, 9(11), 2112.
- Longhi, S., Nijkamp, P., & Poot, J. (2005). A meta-analysis assessment of the effect of immigration on wages. *Journal of Economic Surveys*, 19(3), 451–477.
- López-Arceiz, F. J., Bellosta, A. J., & Rivera, P. (2018). Twenty years of research on the relationship between economic and social performance: A meta-analysis approach. *Social Indicators Research*, 140, 453–484.
- Lu, W., Chau, K. W., Wang, H., & Pan, W. (2014). A decade's debate on the nexus between corporate social and corporate financial performance: A critical review of empirical studies 2002–2011. *Journal of Cleaner Production*, 79, 195–206.
- Lu, W., & Taylor, M. E. (2016). Which factors moderate the relationship between sustainability performance and financial performance? A meta-analysis study. *Journal of International Accounting Research*, 15(1), 1–15.
- Matousek, J., Havranek, T., & Irsova, Z. (2022). Individual discount rates: A meta-analysis of experimental evidence. *Experimental Economics*, 25, 318–358.
- Margolis, J. D., Elfenbein, H. A., & Walsh, J. P. (2009). *Does it pay to be good... and does it matter? A meta-analysis of the relationship between corporate social and financial performance*. Available at SSRN: <https://ssrn.com/abstract=1866371> or <https://doi.org/10.2139/ssrn.1866371>
- Matten, D., & Moon, J. (2008). Implicit" and "Explicit" CSR: A conceptual framework for a comparative understanding of corporate social responsibility. *Academy of Management Review*, 33(2), 404–424.
- Mackinnon, J. G., & Webb, M. D. (2017). Wild bootstrap inference for wildly different cluster sizes. *Journal of Applied Econometrics*, 32(2), 233–254.
- Mcshane, B. B., Gal, D., Gelman, A., Robert, C., & Tackett, J. L. (2019). Abandon statistical significance. *The American Statistician*, 73(s1), 235–245.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097.

- Musso, F., & Risso, M. (2006). CSR within large retailers international supply chains. *Symphonya Emerging Issues in Management*(1), 79–92. (www.unimib.it/symphonya).
- Nelson, J. P., & Kennedy, P. E. (2009). The use (and abuse) of meta-analysis in environmental and natural resource economics: An assessment. *Environmental and Resource Economics*, 42(3), 345–377.
- Oczkowski, E., & Doucouliagos, H. (2015). Wine prices and quality ratings: A meta-regression analysis. *American Journal of Agricultural Economics*, 97(1), 103–121.
- Orlitzky, M. (2001). Does firm size comfound the relationship between corporate social performance and firm financial performance? *Journal of Business Ethics*, 33, 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.
- Penrose, E. T. (1959). *The theory of the growth of the firm*. John Wiley & Sons.
- Printezis, I., Grebitus, C., & Hirsch, S. (2019). The price is right!? A meta-regression analysis on willingness to pay for local food. *Plos One*, 14(5), e0215847.
- Raftery, A. E., Madigan, D., & Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, 92, 179–191.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Sage.
- Ross, S. A. (1973). The economic theory of agency: The principal's problem. *The American Economic Review*, 63(2), 134–139.
- Rusnak, M., Havranek, T., & Horvath, R. (2013). How to solve the price puzzle? A meta-analysis. *Journal of Money, Credit and Banking*, 45(1), 37–70.
- Shahzad, A. M., & Sharfman, M. P. (2017). Corporate social performance and financial performance: Sample-selection issues. *Business & Society*, 56(6), 889–918.
- Stanley, T. D. (2005). Beyond publication bias. *Journal of Economic Surveys*, 19(3), 309–345.
- Stanley, T. D. (2008). Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection. *Oxford Bulletin of Economics and Statistics*, 70(1), 103–127.
- Stanley, T. D., & Doucouliagos, H. (2010). Picture this: A simple graph that reveals much ado about research. *Journal of Economic Surveys*, 24(1), 170–191.
- Stanley, T. D., & Doucouliagos, H. (2012). *Meta-Regression Analysis in Economics and Business. Routledge advances in research methods* (Vol. 5). Routledge.
- Stanley, T. D., Doucouliagos, H., Giles, M., Heckemeyer, J. H., Johnston, R. J., Laroche, P., Nelson, J. P., Paldam, M., Poot, J., Pugh, G., Rosenberger, R. S., & Rost, K. (2013). Meta-analysis of economics research reporting guidelines. *Journal of Economic Surveys*, 27(2), 390–394.
- Stanley, T. D., & Jarrell, S. B. (1989). Meta-regression analysis: A quantitative method of literature surveys. *Journal of Economic Surveys*, 3(2), 161–170.
- Suñé, P., Suñé, J. M., & Montoro, J. B. (2013). Positive outcomes influence the rate and time to publication, but not the impact factor of publications of clinical trial results. *Plos One*, 8(1), e54583.
- Ullmann, A. A. (1985). Data in search of a theory: A critical examination of the relationships among social performance, social disclosure, and economic performance of U.S. firms. *The Academy of Management Review*, 10(3), 540–557.
- Valickova, P., Havranek, T., & Horvath, R. (2015). Financial development and economic growth: A meta-analysis. *Journal of Economic Surveys*, 29(3), 506–526.
- Van Beurden, P., & Gössling, T. (2008). The worth of values—a literature review on the relation between corporate social and financial performance. *Journal of Business Ethics*, 82(2), 407–424.
- Verbeek, M. (2012). *A guide to modern econometrics* (4th ed.). John Wiley & Sons.
- Vishwanathan, P., Van Oosterhout, H. J., Heugens, P. P. M. A. R., Duran, P., & Essen, M. (2020). Strategic CSR: A concept building meta-analysis. *Journal of Management Studies*, 57(2), 314–350.
- Von Arx, U., & Ziegler, A. (2014). The effect of corporate social responsibility on stock performance: New evidence for the USA and Europe. *Quantitative Finance*, 14(6), 977–991.
- Waddock, S. A., & Graves, S. B. (1997). The corporate social performance-financial performance link. *Strategic Management Journal*, 18(4), 303–319.
- Wang, K., & Shailer, G. (2015). Ownership concentration and firm performance in emerging markets: A meta-analysis. *Journal of Economic Surveys*, 29(2), 199–229.

- Wang, Q., Dou, J., & Jia, S. (2016). A meta-analytic review of corporate social responsibility and corporate financial performance. *Business & Society*, 55(8), 1083–1121.
- Wang, Z., & Sarkis, J. (2017). Corporate social responsibility governance, outcomes, and financial performance. *Journal of Cleaner Production*, 162, 1607–1616.
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p -values: Context, process, and purpose. *The American Statistician*, 70(2), 129–133.
- Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a world beyond “ $p < 0.05$ ”. *The American Statistician*, 73(s1), 1–19.
- Weber, S. (2010). bacon: An effective way to detect outliers in multivariate data using Stata and (Mata). *The Stata Journal*, 10(3), 331–338.
- Zeugner, S., & Feldkircher, M. (2015). Bayesian model averaging employing fixed and flexible priors: The BMS package for R. *Journal of Statistical Software*, 68(4), 1–37.
- Zigraiova, D., & Havranek, T. (2016). Bank competition and financial stability: Much ado about nothing? *Journal of Economic Surveys*, 30(5), 944–981.
- Zigraiova, D., Havranek, T., Irsova, Z., & Novak, J. (2021). How puzzling is the forward premium puzzle? A meta-analysis. *European Economic Review*, 134, 103714.

How to cite this article: Hirsch, S., Petersen, T., Koppenber, M., & Hartmann, M. (2023). CSR and firm profitability: Evidence from a meta-regression analysis. *Journal of Economic Surveys*, 37, 993–1032. <https://doi.org/10.1111/joes.12523>

APPENDIX

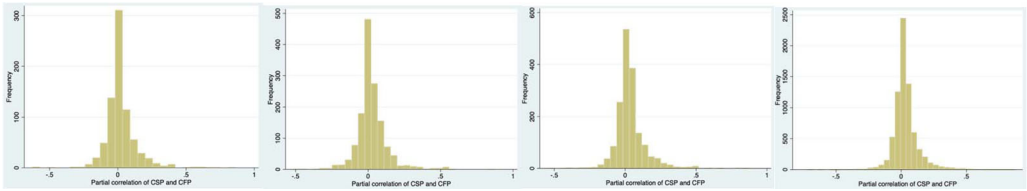


FIGURE A1 Frequency distribution of the partial correlation coefficient ($\hat{\beta}_{CSP}$): subgroups of CSP dimensions and robustness check for precision measure

Notes: $\hat{\beta}_{CSP}$ on the x-axis; the y-axis represents the frequency within the respective interval.

From left to right: only corporate governance CSP dimension considered ($n = 779$); only environmental CSP dimension considered ($n = 1425$); only social CSP dimension considered ($n = 1717$); observations where only asterisks are reported are excluded ($n = 7516$).

[Colour figure can be viewed at wileyonlinelibrary.com]

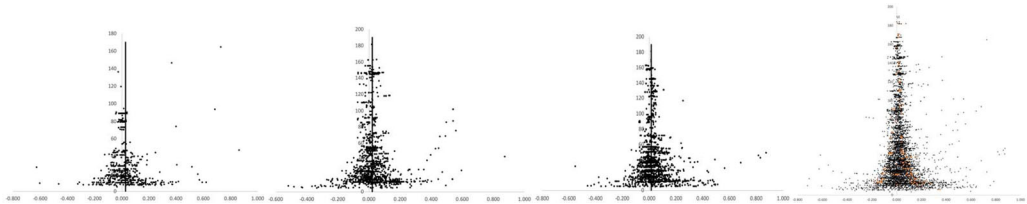


FIGURE A2 Funnel plots for the partial correlation coefficient ($\hat{\beta}_{CSP}$): subgroups of CSP dimensions and robustness check for precision measure

Notes: $\hat{\beta}_{CSP}$ on the x-axis; Precision measured as the inverse standard error on the y-axis. The vertical lines reflect the “true” $\hat{\beta}_{CSP}$ measured as the mean of the 10% most precise values.

From left to right: only corporate governance CSP dimension considered ($n = 779$), true value: 0.027; only environmental CSP dimension considered ($n = 1425$), true value: 0.020; only social CSP dimension considered ($n = 1717$), true value: 0.017; observations where only asterisks are reported are colored in orange.

[Colour figure can be viewed at wileyonlinelibrary.com]

TABLE A1 Overview of previous meta-regression-analyses

Author(s) (publication year)	Journal	Type	Time span	Sample studies/ observations	CSP- dimensions	Number of included study attributes	Finding
Allouche and Laroche (2005)	White Paper	Meta-regression analysis	1972–2003	82/373	All	32	Weak positive correlation
Busch and Friede (2018)	Corp Soc Responsib Environ Manag	Second-order meta-analysis	1997–2015	25 ^a /1214	All	–	Weak positive correlation
Dixon-Fowler et al. (2013)	J Bus Ethics	Meta-analysis	1970–2009	71/71	Environmental	25	Weak positive correlation
Endrikat (2016)	J Bus Ethics	Meta-analysis	1990–2013	32/53	Environmental	11	Weak positive correlation
Friede et al. (2015)	Journal of Sustainable Finance & Investment	Second-order meta-analysis	1990–2014	60 ^b /1678	All	–	Weak positive correlation
Hang et al. (2019)	Bus Strat Env	Meta-regression analysis	1990–(-)	142/893	Environmental	14	Mixed results
Hou et al. (2016)	Asia Pac J Manag	Meta-regression analysis	2001–2015	28/89	Social and environmental	6	Weak positive correlation
Huang (2021)	Account Finance	Second-order meta-analysis	1980–2019	21	All	–	Weak positive correlation
Huang et al. (2020)	International Review of Financial Analysis	Meta-regression analysis	1975–2019	98/437	All	16	Mixed results (positive if controlled for economic fluctuations)

(Continues)

TABLE A1 (Continued)

Author(s) (publication year)	Journal	Type	Time span	Sample studies/ observations	CSP- dimensions	Number of included study attributes	Finding
López-Arceiz et al. (2018)	Soc Indic Res	Meta-regression analysis	1996–2017	83/678	Social	16	Positive correlation
Lu et al. (2014)	Journal of Cleaner Production	Meta-analysis	2002–2011	84/	All	22	Mixed results
Lu, Taylor (2016)	Journal of International Accounting Research	Meta-analysis	–	36/198	Environmental & social	24	Weak positive correlation
Margolis et al. (2009)	SSRN Journal	Meta-analysis	1972–2007	167/192	All	11	Weak positive correlation
Orlitzky (2001)	Journal of Business Ethics	Meta-analysis	1975–1997	20/41	Environmental & social	3	Weak positive correlation
Orlitzky et al. (2003)	Organization Studies	Meta-analysis	1970	52/388	Environmental & social	12	Weak positive correlation
Vishwanathan et al. (2020)	J. Manage. Stud.	Meta-analytic structural equation modeling	1978–2016	334/1446	All	21	Weak positive correlation
Wang et al. (2016)	Business & Society	Meta-analysis	2003–2012	42/119	All	12	Weak positive correlation

^a Meta-studies.

TABLE A2 Descriptive statistics of CSP–CFP meta-data

Variable label	Definition	Mean	sd	Min	Max
<i>Dependent variable</i>					
β_{CSP}	Partial correlation coefficient for the CFP–CSP relationship	0.035	0.107	−0.719	0.875
β_{CSP} only_gov	Partial correlation for the CFP–CSP relationship if only the corporate governance CSP dimension is considered	0.023	0.120	−0.627	0.862
β_{CSP} only_env	Partial correlation for the CFP–CSP relationship if only the environmental CSP dimension is considered	0.024	0.110	−0.518	0.873
β_{CSP} only_soci	Partial correlation for the CFP–CSP relationship if only the social CSP dimension is considered	0.035	0.109	−0.551	0.875
<i>Precision</i>					
SE $_{\beta}$	Standard error of the partial correlation coefficient	0.038	0.034	0.005	0.314
Number of obs	Number of observations in the underlying dataset	4211.214	6224.585	10	36,026
Test statistics	1 if the study reports test statistics for coefficients, 0 if significance levels are only indicated with asterisks	0.964		0	1
<i>Research design</i>					
Journal	1 if the primary regression is from a peer-reviewed journal article; 0 otherwise	0.979		0	1
Non peer-reviewed	1 if the primary regression is from a working or conference paper, book, or thesis; 0 otherwise	0.021		0	1
Impact factor	1 if the primary regression is from a journal with IF at time of publication; 0 otherwise	0.776		0	1
CSP–CFP focus	1 if the paper focuses explicitly on the CSP–CFP relationship	0.581		0	1
Lastyear	Last year of the analyzed period—1989	24.341	4.063	1	31
Industrials	1 if the sample consists of firms from the industrial sector (GICS 20); 0 otherwise	0.030		0	1
Financials	1 if the sample consists of firms from the financial sector (GICS 40); 0 otherwise	0.082		0	1
Consumer goods	1 if the sample consists of firms from the consumer discretionary/staples sector (GICS 25 and 30); 0 otherwise	0.032		0	1
Other industries	1 if the sample consists of firms, which do not operate in the industrial, financial, or consumer goods sector or of a combination of multiple industries; 0 otherwise	0.855		0	1
EU	1 if the sample consists of firms from the European Union only; 0 otherwise	0.148		0	1
U.S.	1 if the sample consists of firms from North America (USA/CAN) only; 0 otherwise	0.338		0	1
Developing	1 if the sample consists of firms from developing countries only; 0 otherwise	0.083		0	1
China	1 if the sample consists of firms from China only; 0 otherwise	0.136		0	1
BRICS	1 if the sample consists of firms from BRICS countries only; 0 otherwise	0.032		0	1

(Continues)

TABLE A2 (Continued)

Variable label	Definition	Mean	sd	Min	Max
Other_countries	1 if the sample consists of firms, which do not have their origins in North America, the EU, China, developing countries, BRICS countries or relate to a combination of these countries; 0 otherwise	0.264		0	1
CSP Index	1 if the CSP variable was created based on a commercial CSP index; 0 otherwise	0.820		0	1
CSP disclosed	1 if the CSP variable was created based on companies' disclosure regarding CSP; 0 otherwise	0.174		0	1
CSP questionnaire	1 if the CSP variable is based on CSP data collected with questionnaires; 0 otherwise	0.006		0	1
CSP binary	1 if the primary CSP variable is a dummy variable, 0 otherwise	0.053		0	1
CSP_squared	1 if CSP is squared, 0 otherwise	0.019		0	1
CSP_lag	1 if CSP is lagged by one or more periods, 0 otherwise	0.175		0	1
CSP_interaction	1 if an interaction of CSP and a control variable is included	0.106		0	1
CFP accounting	1 if CFP is measured with an accounting-based measure, 0 otherwise	0.621		0	1
CFP market	1 if CFP is measured with a market-based measure, 0 otherwise	0.379		0	1
CSP dependent	1 if CSP is the dependent variable in the respective regression, 0 otherwise	0.456		0	1
All_CSP	1 if all CSP dimensions are included in the CSP variable, 0 if otherwise	0.311		0	1
Only_gov_CSP	1 if only the governance dimension is included in the CSP variable, 0 otherwise	0.100		0	1
Only_env_CSP	1 if only the environmental dimension is included in the CSP variable, 0 otherwise	0.183		0	1
Only_soc_CSP	1 if only the social dimension is included in the CSP variable, 0 otherwise	0.220		0	1
Soc&Env_CSP	1 if the social and the environmental dimension are both included in the CSP variable, 0 otherwise	0.163		0	1
Env&Gov_CSP	1 if the governance and the environmental dimension are both included in the CSP variable, 0 otherwise	0.004		0	1
Soc&Gov_CSP	1 if the governance and the social dimension are both included in the CSP variable, 0 otherwise	0.019		0	1
CSP_control	1 if the regression includes more than one CSP variable (Equation 2), 0 otherwise (Equation 1)	0.373		0	1
Size_control	1 if regression controls for the size of the sample firms, 0 otherwise	0.940		0	1
Industry_control	1 if regression controls for the industries of the sample firms, 0 otherwise	0.599		0	1
Risk_control	1 if regression controls for the risk of the sample firms, 0 otherwise	0.663		0	1
R&D_control	1 if regression controls for the R&D spending of the sample firms, 0 otherwise	0.344		0	1

(Continues)

TABLE A2 (Continued)

Variable label	Definition	Mean	sd	Min	Max
Debt_control	1 if regression controls for the debt of the sample firms, 0 otherwise	0.881		0	1
Advertising_control	1 if regression controls for advertising expenditure of the sample firms, 0 otherwise	0.113		0	1
Age_control	1 if regression controls for the age of the sample firms, 0 otherwise	0.249		0	1
Profit_control	1 if regression controls for an additional profit variable, 0 otherwise	0.394		0	1
Share_control	1 if regression controls for shareholder characteristics of the sample firms, 0 otherwise	0.401		0	1
Board_control	1 if regression controls for board characteristics of the sample firms, 0 otherwise	0.340		0	1
Number of years	Length of the observation period measured in number of years	9.197	5.093	1	25
Cross	1 if the estimation is based on cross-sectional data; 0 otherwise	0.076		0	1
Panel	1 if the estimation is based on panel data; 0 otherwise	0.924		0	1
OLS	1 if the regression is estimated with an OLS estimator, 0 otherwise	0.499		0	1
Fixed	1 if the regression is estimated with a fixed effect estimator, 0 otherwise	0.221		0	1
GMM	1 if the regression is estimated with the generalized method of moments, 0 otherwise	0.103		0	1
2SLS	1 if the regression is estimated with a 2 stage least squares model, 0 otherwise	0.067		0	1
Other_estimator	1 if the estimation is based on a different approach, 0 otherwise	0.110		0	1

TABLE A3 Assessment of excess variation in reported coefficients

	Full sample	Only corporate governance CSP dimension	Only environmental CSP dimension	Only social CSP dimension
“True” $\hat{\beta}_{CSP}$	0.023	0.027	0.020	0.017
% $\left \frac{\hat{\beta}_{CSP} - \beta_{CSP}}{se(\hat{\beta}_{CSP})} \right > 1.96$	0.376	0.327	0.331	0.323
z-value ($H_0: \pi = 0.05$)	132.063***	35.517***	48.710***	51.949***
Obs.	7800	779	1425	1717

Note: “True” $\hat{\beta}_{CSP}$ measured as the mean of the 10% most precise values; π is proportion.

*** $p < .01$.

TABLE A4 WLS FAT-PET MRA results for subsamples of individual CSP dimensions and robustness checks for full sample

Variable	Subsamples				Full sample			Focus explicitly on CSR-CFP and CFP-dependent (7)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Constant	0.039 (0.034)	0.020** (0.008)	0.013** (0.006)	0.017*** [0.002]	0.017*** (0.005)	0.01*** (0.004)	0.013*** (0.004)	
se	-0.683 (0.865)	0.170 (0.189)	0.641** (0.273)	0.490*** [0.008]	0.472*** (0.173)	0.658*** (0.166)	0.610*** (0.174)	
Obs.	779	1425	1717	7800	7516	4528	3806	
R ²	0.005	0.001	0.017	0.006	0.005	0.022	0.021	
F-value	0.62	0.81	5.51**		7.47***	15.81***	12.28***	
# of studies	80	207	164	512	480	262	238	

Note: Study-cluster robust standard errors in parentheses; *p*-values in brackets.

****p* < .01, ***p* < .05, **p* < .1.

TABLE A5 Robustness checks MRA results with full set of research design characteristics

Variable label	WLS (SE split)		Wild bootstrap		Only CFP-dependent		Asterisks-only excluded		BMA (no. of obs.)		Post BMA (no. of obs.)		FMA (no. of obs.)	
	Coef.	SE	Coef.	p	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
Constant	0.008	0.029	0.006	0.882	0.084**	0.035	0.005	0.033			0.006	0.016	0.000	7.442
se			0.408**	0.040	0.408*	0.232	0.372*	0.190	4.543	8.723	5.582***	1.654	0.467***	0.169
1/no. of obs.														
$se^*D(\hat{\beta}_{CSP} \geq 0)$	1.759***	0.153												
$se^*D(\hat{\beta}_{CSP} < 0)$	-2.178***	0.211												
Journal	-0.033	0.027	-0.021	0.588	-0.027	0.035	-0.021	0.030	-5.22e-5	0.001	0.012	0.000	0.018	
Impact factor	0.003	0.007	0.004	0.758	0.010	0.008	0.004	0.008	0.001	0.002	0.176	0.000	0.004	
Last year	0.001	0.001	0.001	0.522	9.09e-6	0.001	0.001	0.001	1.92e-7	1.93e-5	0.010	0.000	6.00e-04	
Industrials	0.018*	0.010	0.019*	0.100	0.024	0.018	0.019	0.012	0.003	0.007	0.214	0.005	0.015	
Financials	0.014	0.011	0.015	0.368	0.008	0.015	0.015	0.016	3.12e-4	0.002	0.029	0.000	0.015	
Consumer goods	0.039	0.025	0.048	0.134	0.092	0.054	0.059	0.037	0.044	0.008	1.000	0.040	0.014	0.036
EU	0.003	0.006	0.002	0.764	-0.003	0.014	0.004	0.009	1.39e-4	0.001	0.024	0.000	0.002	
Developing	0.019	0.013	0.011	0.558	-0.028*	0.016	0.016	0.018	2.66e-4	0.002	0.022	0.000	0.012	
China	0.042***	0.015	0.041**	0.022	-0.016	0.017	0.042***	0.016	0.044	0.002	1.000	0.044***	0.015	0.042***
BRICS	-0.003	0.013	-0.015	0.354	-0.033	0.020	-0.008	0.016	-4.23e-4	0.003	0.027	0.000	0.014	
Other_countries	-0.006	0.005	-0.004	0.500	-0.007	0.008	-0.004	0.006	-1.78e-4	0.001	0.050	0.000	0.004	
CSP disclosed	-0.018	0.012	-0.018	0.324	-0.002	0.013	-0.020	0.016	-0.008	0.008	0.507	-0.014	0.012	0.000
CSP questionnaire	0.027	0.019	0.046*	0.070	0.045*	0.026	0.039*	0.023	0.006	0.016	0.152	0.015	0.037	
CSP binary	-0.023***	0.006	-0.033***	0.008	-0.031***	0.011	-0.033***	0.010	-0.031	0.004	1.000	-0.030***	0.009	-0.029***
CSP_squared	-0.020***	0.007	-0.047***	0.002	-0.044***	0.013	-0.044***	0.009	-0.044	0.008	1.000	-0.042***	0.008	-0.033**
CSP_lag	-0.007*	0.004	-0.011**	0.030	-0.005	0.004	-0.011**	0.005	-0.012	0.002	1.000	-0.012***	0.004	-0.004
CSP_interaction	0.001	0.004	-2.76e-4	0.970	0.001	0.004	0.001	0.005	4.81e-6	2.29e-4	0.008	0.000	3.00e-04	

(Continues)

TABLE A 5 (Continued)

Variable label	WLS (SE split)		Wild bootstrap		Only CFP-dependent		Asterisks-only excluded		BMA (no. of obs.)		Post BMA (no. of obs.)		FMA (no. of obs.)	
	(1) Coef.	(1) SE	(2) Coef.	(2) p	(3) Coef.	(3) SE	(4) Coef.	(4) SE	(5) Post mean	(5) Post SD	(6) Coef.	(6) SE	(7) Coef.	(7) SE
CFP accounting	0.003	0.004	0.005	0.294	-0.001	0.004	0.006	0.005	0.005	0.002	0.881	0.006	0.000	0.005
CSP-dependent	-0.011**	0.005	-0.017***	0.010	-	-	-0.017***	0.006	-0.018	0.002	1.000	-0.019***	0.005	-0.010*
Only_gov_CSP	0.006	0.010	-0.003	0.880	-0.013	0.013	-0.003	0.019	0.009	-5.81e-6	3.21e-4		0.000	0.004
Only_env_CSP	0.006	0.009	0.004	0.706	-0.008	0.007	0.005	0.010	1.55e-5	2.57e-4	0.012		0.000	0.004
Only_soc_CSP	0.004	0.005	0.004	0.546	0.002	0.006	0.005	0.006	-3.67e-6	1.79e-4	0.009		0.000	0.004
Soc&Env_CSP	0.001	0.004	0.002	0.702	0.001	0.007	0.002	0.005	7.60e-6	1.74e-4	0.009		0.000	0.002
Env&Gov_CSP	-0.002	0.009	0.012	0.348	0.008	0.014	0.014	0.012	8.44e-5	0.002	0.008		0.000	0.013
Soc&Gov_CSP	0.005	0.007	0.004	0.716	0.006	0.009	0.004	0.010	3.80e-5	0.001	0.011		0.000	0.004
CSP_control	-0.004	0.004	-0.008	0.140	-0.011***	0.004	-0.008*	0.005	-0.007	0.002	0.974	-0.007**	0.004	0.000
Size_control	0.034**	0.015	0.032	0.106	-0.006	0.017	0.033**	0.016	0.034	0.003	1.000	0.033**	0.016	0.030***
Industry_control	0.001	0.004	0.004	0.328	-0.005	0.006	0.004	0.005	0.001	0.002	0.186		0.000	0.005
Risk_control	-0.007	0.005	-0.008	0.254	-0.005	0.006	-0.008	0.006	-0.001	0.002	0.237		0.000	0.008
R&D_control	0.002	0.016	0.003	0.560	-1.23e-4	0.006	0.004	0.006	0.001	0.002	0.249		0.000	0.003
Debt_control	0.002	0.016	0.007	0.788	0.019	0.013	0.006	0.017	-1.78e-6	3.14e-4	0.009		0.000	0.008
Advertising_control	0.001	0.004	0.003	0.544	-0.002	0.006	0.003	0.005	5.91e-5	4.89e-4	0.022		0.000	0.003
Age_control	4.37e-4	0.006	0.004	0.602	0.004	0.006	0.004	0.006	0.001	0.002	0.248		0.000	0.004
Profit_control	-0.007	0.005	-0.010*	0.092	-0.009**	0.004	-0.011*	0.006	-0.009	0.002	1.000	-0.008	0.005	-0.007**

(Continues)

TABLE A 5 (Continued)

Variable label	WLS (SE split)		Wild bootstrap		Only CFP-dependent		Asterisks-only		BMA (no. of obs.)		Post BMA (no. of obs.)		FMA (no. of obs.)		
	Coef.	SE	Coef.	p	Coef.	SE	Coef.	SE	Post mean	Post SD	PIP	Coef.	SE	Coef.	SE
Share_control	0.008*	0.005	0.007	0.212	0.001	0.008	0.007	0.005	0.008	0.002	0.999	0.008*	0.005	0.003	0.006
Board_control	-2.81e-4	0.005	0.002	0.744	0.007	0.007	0.001	0.006	7.91e-5	0.001	0.030			0.000	0.002
Number of years	-3.83e-4	4.99e-4	-4.45e-4	0.452	-0.002***	0.001	-4.19e-4	0.001	-5.80e-6	4.29e-5	0.026			0.000	4.00e-04
Cross	-0.010	0.010	-0.011	0.346	-0.010	0.015	-0.011	0.012	-6.24e-5	0.001	0.010			0.000	0.011
Fixed	-0.015***	0.005	-0.020***	0.004	-0.022***	0.006	-0.020***	0.005	-0.021	0.002	1.000	-0.021***	0.005	-0.021***	0.005
GMM	-0.013***	0.005	-0.022***	0.006	-0.018**	0.008	-0.023***	0.008	-0.020	0.003	1.000	-0.019***	0.005	-0.017**	0.008
2SLS	-0.016***	0.006	-0.014**	0.040	-0.011**	0.005	-0.014**	0.006	-0.013	0.002	1.000	-0.013**	0.005	-0.011***	0.004
Other_estimator	-0.004	0.006	-0.009	0.226	-0.009	0.009	-0.009	0.008	-0.005	0.005	0.477			0.000	0.009
Obs.	7800		7800		4246		7516		7800			7800			7800
R ²	0.321		0.134		0.167		0.135					0.125			
F-value	24.53***				6.63***		6.43***					6.59***			
Studies	512		512		295		480		512			512			512

Note: Study-cluster robust standard errors in parentheses. Base group: non peer-reviewed, mixed industries, US, CSP-Index, CFP-market, All_CSP, cross, OLS. PIP = "posterior inclusion probability"; post-BMA regressions are performed using variables with PIP > 0.5. ***p < .01, **p < .05, *p < .1.