Strategic CSR: A Concept Building Meta-Analysis

Pushpika Vishwanathan, Hans (J.) van Oosterhout, Pursey P. M. A. R. Heugens, Patricio Duran and Marc van Essen

University of Amsterdam; Erasmus University; Universidad Adolfo Ibáñez; University of South Carolina

ABSTRACT

This study develops the concept of Strategic Corporate Social Responsibility (Strategic CSR) by meta-analyzing the available empirical evidence on the relationship between CSR and corporate financial performance (CFP). Using meta-analytic structural equation modeling on effect size data from 344 primary studies, our study documents four empirical mechanisms explaining how CSR positively affects CFP: 1) enhancing firm reputation, 2) increasing stakeholder reciprocation, 3) mitigating firm risk, and 4) strengthening innovation capacity. We propose these four mechanisms to identify four causally relevant attributes that allow us to conceptually distinguish Strategic CSR from CSR more generally. Our findings indicate that the four mechanisms combined explain 20 per cent of the CSR-CFP relationship, suggesting that considerable room remains for future empirical research. The development of an empirically informed, causal conceptualization of Strategic CSR responds to a long-heard call for better-specified concepts in empirical CSR research.

Keywords: concept formation, corporate social responsibility, financial performance, meta-analysis, strategy

INTRODUCTION

The concept of corporate social responsibility (CSR) has inspired empirical research in management for almost half a century (Aguinis and Glavas, 2012). Early empirical CSR research initially had the (often only implicit) ambition of showing that the socially beneficial activities denoted by the CSR concept could be strategically justified by...
their positive performance implications for the firms undertaking them (Rowley and Berman, 2000; Wood and Jones, 1995). Orlitzky and colleagues (2003) synthesized the findings of 52 primary studies on the relationship between CSR and corporate financial performance (CFP) produced in these early years. By documenting a small positive association between CSR and CFP, they corroborated that CSR can indeed be justified strategically.

Over the last two decades, however, CSR research has shifted its focus from investigating the CSR-CFP relationship as a whole to identifying the concrete empirical mechanisms through which CSR activities affect firm-level financial outcomes. This shift seems to have been motivated by three reasons. First, the documentation of a small positive relationship between CSR and CFP by Orlitzky and colleagues (2003) obviated the justificatory ambition to show that CSR mattered, thereby making room for the explanatory question of how CSR mattered instead (Barnett, 2007; Wang et al., 2016). Second, scholars became increasingly critical of the coarse-grained and over-inclusive approach to conceptualizing CSR, which produced a concept referring to a diverse set of empirical phenomena (Gond and Crane, 2010). As a response, finer-grained research was increasingly undertaken, focusing on specific types of CSR activities and their indirect effect on firm performance (Flammer, 2013; King and Lenox, 2002; Wang and Qian, 2011). Third, CSR research became methodologically more sophisticated as scholars started to use more advanced research designs (Flammer, 2013, 2015a; Ortiz-de-Mandojana and Bansal, 2016), better data (Henisz et al., 2014; Lev et al., 2010; Zhao and Murrell, 2016), and more rigorous analytical techniques (Cheng et al., 2014; Surroca et al., 2010), that required a shift towards more focused research questions (Godfrey et al., 2009; Koh et al., 2014; Ortiz-de-Mandojana and Bansal, 2016).

As individual CSR-CFP studies became finer-grained and more focused, however, the field as a whole became more fragmented. A diverse and growing body of empirical findings has emerged identifying a wide variety of empirical mechanisms linking CSR to CFP. One of the most persistent conceptual critiques on CSR involved its lack of theoretical foundations (Ullmann, 1985), resulting in a poorly specified construct (Rowley and Berman, 2000; Wood and Jones, 1995) that has highly permeable boundaries (Lockett et al., 2006). This has led some scholars to claim that CSR hardly meets even the most basic requirements for explanatory concepts in social science (Van Oosterhout and Heugens, 2008). Although several attempts at conceptual clarification were made (Clarkson, 1995; Swanson, 1995; Windsor, 2006), scholars seem to have accepted that CSR is better conceived of as an umbrella term rather than a well-defined theoretical concept (Gond and Crane, 2010). But as empirical research on CSR and on the CSR-CFP relationship continues to proliferate, the absence of a well-defined theoretical concept increasingly hinders the development of the field (Lockett et al., 2006; McWilliams et al., 2006; Pfeffer, 1993), since a minimal degree of consensus regarding the defining features of a concept is required for scholars to effectively engage with each other to advance common knowledge (Kuhn, 1962; Suddaby, 2010).

Within the general field of CSR research, we focus specifically on the stream of research investigating the CSR-CFP relationship in order to isolate and develop the concept of ‘Strategic CSR’. Following Goertz’ (2006) causal approach to explanatory concepts in social science, we develop a three-level conceptual structure of the concept of Strategic
CSR in an effort to identify its causally most relevant properties for explaining financial performance differentials between firms. We review the empirical CSR-CFP literature, and ask three questions about the current state of research. First, what do we currently know about how CSR affects firm financial outcomes? Specifically, what are the most important empirical mechanisms through which the variety of CSR activities investigated in the literature affect CFP? Second, does our current knowledge about how CSR activities affect CFP ‘exhaust’ the overall CSR-CFP relationship, that is, is the relationship between CSR and CFP fully explained (i.e. fully mediated) by the four empirical mechanisms that we identify in this study? Third, based on the findings of this study, what potentially fruitful avenues can we suggest for future research on (Strategic) CSR?

We answer these questions using advanced meta-analytic techniques on the empirical evidence accumulated over five decades of empirical CSR-CFP research. Based on a review of the literature, we identify four theoretical mechanisms for which empirical evidence is sufficiently available. We then use meta-analytical structural equation modelling (MASEM: Bergh et al., 2016) on a combined sample of 402,863 firm-year observations retrieved from 344 empirical studies to provide a simultaneous meta-analytic test of the four empirical mechanisms identified.

We seek to make three contributions to the CSR literature. First, we synthesize the growing and increasingly fragmented body of research on the relationship between CSR and CFP. We document four mechanisms through which CSR activities may contribute to the financial bottom line of firms: firm reputation, stakeholder reciprocation, risk mitigation, and innovation capacity. Second, we rely on these four mechanisms to theoretically develop the concept of Strategic CSR by defining it in terms of its causally most relevant attributes in explaining CFP. By developing an evidence-based conceptualization of Strategic CSR that can be isolated and carved out from the conceptual domain of CSR more generally, we respond to the long-heard call for better-specified concepts in CSR research around which CSR researchers may unite and contribute to the development of a shared paradigm (Gond and Crane, 2010; Jones, 1995; Lockett et al., 2006). Third, our findings serve as a guide for future CSR-CFP research, both in terms of establishing what we already know and in identifying possible avenues for innovative research contributions still to be made.

THEORETICAL BACKGROUND

The concept of CSR initially developed in the context of a normative debate, as much of the early CSR literature evolved around the normative question whether firms have an obligation to promote social betterment over and above their economic and legal obligations (Carroll, 1999; Frederick, 1994; Matten et al., 2003). In this debate, CSR was mostly residually conceptualized as socially valuable firm activities not required by law or shareholder interests (McWilliams and Siegel, 2001), which explains the coarse and over-inclusive meaning that the concept subsequently acquired in the literature (Van Oosterhout and Heugens, 2008). The move to position CSR as a positive explanatory concept emerged from one particular line of argument in this debate; that CSR activities can be strategically justified because, next to promoting socially beneficial
outcomes, they will also positively affect CFP (Rowley and Berman, 2000; Wood and Jones, 1995). In support of this argument, an entire stream of empirical CSR-CFP research emerged. Since its inception, however, the concept of CSR has met with various critiques. Referred to as a(n) ‘contestable’ (Windsor, 2006, p. 93), ‘elusive’ (Clarkson, 1995, p. 92), and even ‘epiphenomenal’ concept (Van Oosterhout and Heugens, 2008, p. 210), the main concern has always been that CSR is theoretically poorly defined (Rowley and Berman, 2000; Ullmann, 1985; Wood and Jones, 1995) and therefore does not satisfy even the most basic requirements for explanatory concepts in social science.

A causal approach to conceptualizing Strategic CSR

To carve out and further develop the explanatory concept of Strategic CSR from the conceptual domain denoted by the concept of CSR more generally, we follow Goertz’ (2006) causal approach to concept formation. Arguing that explanatory concepts in social science are essentially constructed from our causal knowledge on how the phenomenon denoted by the concept interacts with other phenomena, Goertz proposes to define explanatory social science concepts in terms of their causally most relevant general attributes. This causal approach is also a positivist approach to concept formation, because the causally most relevant attributes that jointly define a concept can only be identified through empirical research on the mechanisms through which the phenomenon interacts with other phenomena (a process sometimes referred to as ‘nomological validation’ (Cronbach and Meehl, 1955)). Empirically validated causal propositions that prior research has found to explain social reality are therefore deeply embedded in explanatory social science concepts, and can be used to foster an ever better theoretical understanding of these concepts themselves.1

As the bulk of empirical research has used the CSR concept to explain CFP, we are able to develop an understanding of Strategic CSR by identifying those general attributes of CSR that prior research has found to be causally most relevant in explaining CFP. The concept of Strategic CSR was first coined by Baron (2001), and subsequently developed further by McWilliams and Siegel (2001) and Waldman and colleagues (2006), who explicitly connected it to the strategic management field’s academic objective of explaining firm competitive advantages from observed performance differentials between firms (Nelson, 1991; Rumelt et al., 1991). The development of strategic management concepts by inferring them from performance differentials between firms is an established conceptualization strategy in the strategic management literature and has been foundational for some of the field’s core theories and concepts (for example, the resource-based view: Barney, 1991, 2001; Priem and Butler, 2001; and dynamic capabilities: Winter, 2003; Zahra et al., 2006).2

It is also a useful strategy to conceptually demarcate and carve out Strategic CSR from the empirical domain denoted by the CSR concept more generally, as there may also exist CSR activities that do not enhance CFP (Wright and Ferris, 1997). Our conceptualization strategy to demarcate Strategic CSR from CSR more generally is graphically illustrated by the Venn diagram in Figure 1. The outer rectangle captures the universe of all firm activities. Within this universe, CSR activities are those firm activities that ‘appear to further some social good’ (McWilliams and Siegel, 2001, p. 117). Within the set of CSR activities more generally, there are some CSR activities that positively contribute to firm performance and some that do not. Only the former are part of the empirical
domain denoted by the concept of Strategic CSR. We aim to further develop the concept of Strategic CSR by unveiling the mechanisms through which CSR contributes to CFP, and which identify the causally most relevant attributes of Strategic CSR.

To identify the general attributes of Strategic CSR, we survey five decades of empirical research on the CSR–CFP relationship. While early empirical research investigated the direct relationship between CSR and CFP (Aupperle et al., 1985; Cochran and Wood, 1984; Mcguire et al., 1988), more recent research has documented that the relationship is mediated by other strategically relevant variables (Surroca et al., 2010). By surveying the available research, we seek to uncover the most important mediating mechanisms through which CSR has been found to affect CFP. We focus only on those mechanisms for which sufficient primary studies are available to test the hypotheses developed through our review using MASEM (Bergh et al., 2016). This leads us to identify four empirical mechanisms through which CSR has been found to affect CFP, that is, through: 1) enhanced firm reputation, 2) stakeholder reciprocation, 3) risk mitigation, and 4) improved innovation capacity.

Below, we review the literature on these four mechanisms, and conclude each subsection with articulating a hypothesis that theoretically predicts how CSR affects CFP through this mechanism. We propose that these four mechanisms are conceptually and empirically distinct and identify four different causally relevant general attributes of Strategic CSR. Because our survey suggests that these four mechanisms are the main mechanisms through which CSR has been documented to affect CFP, we also conjecture that they jointly identify the causally relevant general attributes of the Strategic CSR concept exhaustively. We test this conjecture by introducing a fifth hypothesis predicting that the CSR-CFP relationship is fully mediated by these four mechanisms.

**HYPOTHESES DEVELOPMENT**

**The enhanced firm reputation mechanism**

One well-researched empirical mechanism through which CSR has been found to affect CFP involves reputation enhancement. This mechanism captures the strategic benefits
of CSR that result from stakeholders perceiving the firm as a more attractive partner to do business with (Boyd et al., 2010; Fombrun and Shanley, 1990). Turban and Greening (1997) were among the first to demonstrate this effect empirically, showing that a strong CSR rating increases firm attractiveness to prospective employees. Subsequent research corroborated this finding by uncovering that a firm’s involvement in CSR activities signals organizational norms and values, which influences prospective employees’ perceptions of the working conditions at the firm (Backhaus et al., 2002; Greening and Turban, 2000), and increases their willingness to be associated with the firm (Jones et al., 2014).

Firms signal their CSR activities to stakeholders through their advertising or via external ‘infomediaries’ (Carter, 2006; Deephouse and Heugens, 2009; McWilliams and Siegel, 2001; Schuler and Cording, 2006). A CSR activity for which the effect on reputation is particularly well established, for example, is philanthropic donations (Brammer and Millington, 2005; Fombrun and Shanley, 1990; Lev et al., 2010; Williams and Barrett, 2000). Such CSR activities are strategically useful because they attract attention and evoke positive attributions by stakeholders (Godfrey, 2005; Groza et al., 2011). Firms are sometimes criticized for using CSR to convey a positive image without making material changes within the firm (Haack et al., 2012; Marquis and Qian, 2014; Wickert et al., 2016). For example, firms may use corporate philanthropy symbolically, to divert attention away from potential misconduct or to buy goodwill after being accused of misconduct (Du, 2015; Koehn and Ueng, 2010). For the reputation enhancement mechanism to be enacted, however, all that is required is that CSR activities target a broad audience and are visible to both existing and prospective stakeholders.

Research has also shown that CSR reputations may enhance CFP. Because many stakeholders are more attracted to CSR firms, these firms have access to a larger pool of stakeholders with whom they can develop productive relationships (Greening and Turban, 2000). Customers, for example, are more attracted to and derive more satisfaction from purchasing products or services from CSR firms (Luo and Bhattacharya, 2006; Sen and Bhattacharya, 2001), which increases their purchasing intention (Sen et al., 2006), as well as their willingness to pay premium prices (Homburg et al., 2005; Marin et al., 2012). Investors have also been shown to be more attracted to CSR firms. While investment analysts used to give negative investment recommendations for firms with high CSR ratings in the early nineties, they often perceive CSR as a legitimate and even positive signal of a firm’s future profitability today (Ioannou and Serafeim, 2015). Indeed, share prices often increase after the public announcement of CSR initiatives (Arya and Zhang, 2009; Ramchander et al., 2012). We therefore hypothesize:

Hypothesis 1: The CSR–CFP relationship is mediated by firm reputation: CSR is positively related to firm reputation, and firm reputation is positively related to CFP.

The stakeholder reciprocation mechanism

A second mechanism involves stakeholder reciprocation. When firms engage in CSR, they take actions that benefit at least certain stakeholder groups (McWilliams and Siegel, 2001). Stakeholders reciprocate by endorsing the firm, resulting in more cooperative,
productive, and enduring relationships (Bosse and Coughlan, 2016; Bosse et al., 2009; Jones, 1995). In contrast to the reputation enhancement mechanism, for this path to operate the adopted CSR activities must specifically benefit the firm’s existing stakeholders, but need not be visible externally.

Research has documented various ways in which CSR activities create value for stakeholders while also increasing CFP. Employees of CSR firms benefit from practices such as fair pay, a safer work environment, and professional development opportunities (El Akremi et al., 2018). In turn, they experience greater job satisfaction (De Roeck et al., 2016), exude higher levels of organizational commitment (Ali et al., 2010; Brammer et al., 2007), and engage more in organizational citizenship behaviors (Bode et al., 2015; Hansen et al., 2011). Firms profit from such motivated employees, not only because they are more productive, but also because firms can contract with them based on trust, which substantially decreases contracting costs (Jones, 1995).

A firm’s suppliers of financial capital also reciprocate CSR activities. Research has found that high CSR firms tend to engage in elaborate voluntary reporting (Dhaliwal et al., 2011), which reduces information asymmetry and agency costs (Cheng et al., 2014). As a result, debtholders, institutional investors, and other equity providers are willing to offer their resources to the firm at more favourable terms (Cheng et al., 2014; Dhaliwal et al., 2011; El Ghoul et al., 2011).

Local communities and government bodies may similarly endorse the firm in CFP-enhancing ways (Frooman, 1999). Governments depend on a public mandate, which incentivizes them to scrutinize the firms in their jurisdictions for how they handle and affect community interests (Doh and Guay, 2006). Higher levels of community endorsement resulting from CSR may thus lead to more favourable regulatory and enforcement conditions for the firm (Campbell, 2007; Sharma and Henriques, 2005) and higher levels of public procurement (Flammer, 2018; den Hond et al., 2014). Communities may also reciprocate by granting CSR firms a societal license to operate (Henisz et al., 2014; Prno and Slocombe, 2012). For example, firms in extraction industries often sign community benefit agreements, which are contracts between the firm and local communities stating how firms compensate communities for the social and environmental disruptions they will cause (Dorobantu and Odziemkowska, 2017). In the absence of such initiatives, firms may face persistent stakeholder conflicts resulting in a depreciation of intangible assets and investor scepticism (Dorobantu and Odziemkowska, 2017; Henisz et al., 2014). CSR may thus have beneficial effects on a firm’s existing stakeholder relationships, leading to stakeholder reciprocation and higher levels of CFP. Hence:

**Hypothesis 2**: The CSR–CFP relationship is mediated by stakeholder reciprocation: CSR is positively related to stakeholder reciprocation, and stakeholder reciprocation is positively related to CFP.

**The risk mitigation mechanism**

CSR firms are involved with a broader set of issues and engage with more diverse stakeholder groups than firms focusing strictly on their core business and operations (Hart
and Sharma, 2004). With this broader perspective, firms can access new information that can be used to reduce firm-specific risk. Indeed, a prior meta-analysis has found a negative relationship between CSR and firm risk (Orlitzky and Benjamin, 2001). More recent research supports this conclusion (Bansal and Clelland, 2004; Lee and Faff, 2009; Luo and Bhattacharya, 2009; Oikonomou et al., 2012), showing that stock price crash risk (Kim et al., 2014) and firm default risk (Sun and Cui, 2014; Verwijmeren and Derwall, 2010) are both reduced by CSR activities.

There are two main explanations of how CSR can reduce risk. CSR may directly mitigate risk because many CSR activities are especially designed to avoid harm to stakeholders. Such activities are for instance, pollution prevention practices, employee health and safety programs, and fair-trade policies. Implementing these initiatives can reduce the firm’s risk exposure. ISO 14001 certification, for example, requires that firms have clear systems in place for monitoring and measuring environmental performance, creating awareness and competence among all employees, and organizing for emergency preparedness and response (Morrow and Rondinelli, 2002).

CSR may also indirectly reduce firm risk. Through CSR, firms can develop closer relationships with their stakeholders, which makes them more willing to share information with these firms (Harrison et al., 2010). CSR can thus prove instrumental for sensing changes or threats, as stakeholders may transmit early-warning signs to managers (Ortiz-de-Mandojana and Bansal, 2016). Volkswagen’s emission-scandal, for example, was partly caused by its authoritative corporate culture, which made employees afraid of questioning targets or sharing bad news with their superiors (Cremer and Bergin, 2015). Had Volkswagen fostered open and collaborative relationships with its employees, these problems might have been resolved before materializing into a costly public scandal. Thus, stakeholder relationships enable firms to anticipate and prevent foreseeable risks (Klassen and Vereecke, 2012; Ortiz-de-Mandojana and Bansal, 2016; Porter and Kramer, 2007).

Lower firm-specific risk creates value not because it directly generates higher CFP, but because it helps to preserve it (Choi and Wang, 2009; Ortiz-de-Mandojana and Bansal, 2016). Being implied in misconduct exposes firms to direct legal expenses such as litigation costs, administrative fines, and criminal sanctions, but also to indirect expenses such as turnover of key personnel, refinancing costs, and forgone sales (Karpoff and Lott, 1993; Koh et al., 2014). Because CSR firms are able to identify threats early on and correct potentially illegal acts, they can prevent these costs (Mcguire et al., 1988; Orlitzky and Benjamin, 2001). In sum, firms that engage in CSR are more risk-aware, allowing them to manage and mitigate these risks more effectively. Lower firm-specific risk, in turn, reduces the likelihood that firms incur unproductive costs that burden CFP. Hence:

Hypothesis 3: The CSR–CFP relationship is mediated by firm risk: CSR is negatively related to firm risk, and firm risk is negatively related to CFP.

The improved innovation capacity mechanism

Scholars have long recognized the potential of CSR to be a lever of innovation and competitive differentiation (Husted and Salazar, 2006; McWilliams and Siegel, 2001). Because firms that engage in CSR tend to adopt a broader perspective and develop...
closer relationships with stakeholders, they are better able to identify new opportunities (Harrison et al., 2010; Tantalo and Priem, 2016). Close relationships with external stakeholders such as local communities, customers, and environmental groups for example, offer new knowledge pools that can become an important source of innovation (Buysse and Verbeke, 2003; Hart, 1995; Hart and Dowell, 2011). Good relationships with internal stakeholders are also important for innovation. Employees may be more willing to share information with the firm (Aragón-Correa et al., 2013; Harrison et al., 2010), and be better able to overcome the sort of short-term orientation that tends to impede innovation (Flammer and Kacperczyk, 2016; Wang and Bansal, 2012).

Another reason why CSR may enhance the firm's innovativeness is that the adoption of many CSR activities requires the development of existing innovation capabilities or triggers the creation of new capabilities. The extent of a firm's absorptive capacity (Delmas et al., 2011; Luo and Du, 2015) and presence of complementary assets (Christmann, 2000), for example, facilitate the successful implementation of CSR. Researchers have established that firms adopting proactive environmental strategies tend to engage in higher-order learning and continuous innovation (Sharma and Vredenburg, 1998), which may result in the development of new dynamic capabilities (Aragón-Correa and Sharma, 2003). Danone, a food products multinational, offers an illustrative example. Since the early 2000s the firm has been launching base of the pyramid projects in rural areas of developing countries. These initiatives were an important source of strategic renewal, as it encouraged experimentation and learning within Danone, and allowed it to develop unique expertise in creating low-cost solutions (Faivre-Tavignot and Dalsace, 2014).

The relationships between innovation, absorptive capacity, and CFP are well established in the strategy literature (Tsai, 2001; Zahra and George, 2002), but there is increasing evidence for the presence of these relationships in the CSR context as well (Delmas et al., 2011; Tang et al., 2012). Innovations resulting from a firm’s engagement with stakeholders can enhance CFP in at least three ways. First, CSR offers a powerful means of product differentiation (Flammer, 2015b; Hull and Rothenberg, 2008; McWilliams and Siegel, 2001; Siegel and Vitaliano, 2007). Second, radical process innovations, especially in the context of environmental management, may reduce both waste and production costs (Christmann, 2000; Dixon-Fowler et al., 2013; King and Lenox, 2002; Klassen and Whybark, 1999). Finally, as illustrated by the Danone example, CSR initiatives may lead to new business models and offer completely new sources of value creation (Hart and Christensen, 2002; Hart and Sharma, 2004). Therefore:

Hypothesis 4: The CSR–CFP relationship is mediated by innovation capacity: CSR is positively related to innovation capacity, and innovation capacity is positively related to CFP

Does our current knowledge allow an exhaustive conceptualization of Strategic CSR?

Thus far, we have theoretically unveiled and articulated four broad empirical mechanisms through which prior empirical research has suggested CSR to improve CFP. We
propose that these four mechanisms identify the four causally relevant properties of Strategic CSR. In developing explanatory concepts in social science, it is not only important to incorporate the available knowledge about a phenomenon into its conceptual understanding, but also to establish whether this understanding is more or less complete in terms of accounting for the available empirical evidence. Methodologically, this implies assessing whether the four empirical mechanisms identified thus far fully mediate the relationship between CSR and CFP. If not, the implication is that there is at least one alternative causal path connecting CSR and CFP, which is currently absent from our theorizing. To determine whether the current conceptualization of Strategic CSR is indeed exhaustive, we therefore tentatively hypothesize that the CSR-CFP relationship is fully mediated by the four empirical mechanisms identified above:

**Hypothesis 5:** The CSR-CFP relationship is fully mediated by firm reputation, stakeholder reciprocation, firm risk, and innovation capacity

**METHODS**

**Sample and coding**

We used five search strategies to identify CSR-CFP studies. First, we read review articles (Aguinis and Glavas, 2012; McWilliams et al., 2006) and prior meta-analyses (Margolis et al., 2007; Orlitzky, 2015). Second, we examined six electronic databases: (1) ABI/INFORM Global, (2) Business Source Premier, (3) JSTOR, (4) Google Scholar, (5) ECONLIT, and (6) SSRN, using search terms such as: ‘corporate social responsibility’, ‘corporate social performance’, ‘ethical investment’, and ‘green investment fund’. Third, we manually searched 15 leading journals in management and finance, including: Academy of Management Journal, Journal of Corporate Finance, Journal of Management Studies, and Strategic Management Journal. Fourth, we emailed all researchers who had written empirical papers on CSR and CFP with missing effect-size information, asking them for a correlation table and any studies we could not retrieve by other means. Finally, using a two-way ‘snowballing’ technique, we backward-traced all references reported in the identified articles and forward-traced all articles that cited the original articles via Google Scholar. These five search strategies yielded a final sample of 344 primary studies. Of these, 296 were published journal articles and 48 were working papers at the time of our analysis. The publication window ranged from 1978 to 2016.

We then developed a coding protocol (Lipsey and Wilson, 2001) for extracting effect size and sample size information for all the variables in our study, including dependent, independent, mediator, and control variables. One author coded all effect sizes. To assess inter-rater agreement, a second rater coded a subsample of 200 randomly selected effect sizes, after which we computed a chance agreement-corrected measure of inter-rater reliability (Cohen’s kappa coefficient; Cohen, 1960). With a value of 0.90, kappa signified high inter-rater agreement (see Klier et al., 2017; Van Essen et al., 2015).
Measurement

Table I provides a description of all our variables. Existing studies have used many different measures for CSR, including both composite indices and single-item measures such as emission reductions (e.g., King and Lenox, 2002) or philanthropic donations (e.g., Zhang et al., 2010). We included all CSR measures in our study. CFP is operationalized using accounting-based measures such as Return on Equity (ROE) or Return on Assets (ROA), or market-based measures such as Tobin’s Q or market-to-book ratio. To measure firm reputation we included both third-party assessments such as expert ratings and firm expenditures on reputation building (Fombrun and Shanley, 1990). Stakeholder reciprocation has not been measured before in primary studies, so we constructed this measure by capturing any manifestation of positive reciprocal behaviour by stakeholders (Bosse and Phillips, 2016), such as employee commitment, customer satisfaction, and the involvement of government, institutional, or inside investors in the firm’s ownership structure. Firm risk was typically measured as variability in financial performance (e.g., Jensen’s beta or the standard deviation of stock returns). To capture innovation capacity we used both generic measures, such as the fraction of R&D expenses over sales, as well as more CSR-specific ones, such as the degree of environmental innovativeness.

We include several control variables in our analysis. We include prior CFP, since research has shown that this variable affects CSR through the mechanism of organizational slack, as well as current CFP through momentum effects (Waddock and Graves, 1997). We control for firm size, as larger firms tend to be more prone to invest in CSR and profit from it because of the economies of scale involved in the acquisition and deployment of CSR-related resources (McWilliams and Siegel, 2001). We include firm leverage because prior CSR research has shown that resource availability is an important predictor of firms’ ability to profit from CSR (Brammer and Millington, 2008). Finally, we control for capital intensity, as prior research has suggested that incorporating CSR initiatives can be more difficult and more costly to implement in capital-intensive industries (Surroca et al., 2010). All variable operationalizations are listed in Table I.

A common problem in meta-analysis is that the measures used in primary studies tend to involve measurement error. In our case, especially the constructs firm reputation, stakeholder reciprocation, and innovation capacity are often measured through self-report surveys (Boyd, Gove, and Hitt, 2005), which are often not corroborated with archival measurements (Bergh et al., 2016). Following best-practice recommendations (Bergh et al., 2016), we deal with measurement error by applying a conservative 0.80 measurement reliability standard to all hypothesized mediator variables.

Meta-analytic procedure

We used Hunter and Schmidt (2004) meta-analysis (HSMA) to compute the meta-analytic mean association ($\bar{r}_c$) between CSR and CFP, standard deviations, percentage of variance due to sampling error, and 95 per cent Credibility intervals around mean correlations (Whitener, 1990). Since our constructs of interest are usually reported as a continuum, we provide correlations ($\bar{r}_c$) rather than standardized differences (d) (O’Boyle et al., 2012). HSMA ‘allows for the correction of statistical artifact and thus provides a relatively accurate estimate of the true average strength and variance of the
<table>
<thead>
<tr>
<th>Variables</th>
<th>Definition</th>
<th>Operationalizations</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSR</td>
<td>Those firm activities that ‘appear to further some social good’ (McWilliams and Siegel, 2001, p. 117)</td>
<td>(1) Third-party evaluations of the firm’s CSR, including assessments by parties such as MSCI, Fortune Magazine, and stock markets such as Nikkei and FTSE (Brammer et al., 2009a; Kang, 2013); (2) archival measures of actual CSR activities, such as toxic release reports and philanthropic donations (Konar and Cohen, 2001; Zhang et al., 2010); (3) self-reported assessments of CSR in surveys (Chen and Wang, 2011).</td>
</tr>
<tr>
<td>CFP</td>
<td>A firm’s overall financial health over a given period of time.</td>
<td>(1) Accounting-based measures of profitability such as ROE, ROA, and ROI (Strike et al., 2006); (2) Market-based measures of stock returns such as Tobin’s Q or market-to-book ratio (Makni et al., 2009).</td>
</tr>
<tr>
<td>Firm reputation</td>
<td>A perceptual representation of a firm’s past actions and future prospects that describe the firm’s overall appeal to its stakeholders (Fombrun, 1996).</td>
<td>(1) External assessments of firm reputation such as the propensity of favorable press articles, expert ratings of firm reputation, and rankings such as Fortune Magazine (Brammer and Pavelin, 2008); (2) expenditures on reputation building efforts such as advertising expenditures/sales, number of public affairs staff, and number of corporate PR communications of CSR activities (Brammer et al., 2009b).</td>
</tr>
<tr>
<td>Stakeholder reciprocation</td>
<td>Any manifestation of active support to the firm by any of the firm’s existing stakeholder groups.</td>
<td>(1) Employees: employee commitment (Marín et al., 2012); (2) government: government ownership or regulatory supportiveness (Gao, 2008); (3) management: inside ownership (Said et al., 2009); (4) customers: customer satisfaction or customer loyalty (Luo and Bhattacharya, 2006), (5) long-term oriented shareholders: institutional ownership (Mahoney and Roberts, 2007), (6) investors and creditors: cost of capital (El Ghoul et al., 2011).</td>
</tr>
<tr>
<td>Firm risk</td>
<td>Variability in the firm’s financial performance associated with events that primarily affect only that firm (Bansal and Clelland, 2002).</td>
<td>(1) Jensen’s beta, standard deviation of stock returns, standard deviation of ROE over the past years (Jo and Na, 2012).</td>
</tr>
<tr>
<td>Innovation capacity</td>
<td>The firm’s ability to generate, accept, and implement new ideas, processes, products or services (Thompson, 1965).</td>
<td>(1) General measures of innovation capacity such as R&amp;D expenses/sales, age of assets, and innovativeness relative to competitors (Surroca et al., 2010); (2) CSR-specific measures of innovation capacity such as environmental innovativeness, and implementation of novel health and safety systems (Gallego-Alvarez et al., 2011).</td>
</tr>
<tr>
<td>Variables</td>
<td>Definition</td>
<td>Operationalizations</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Prior CFP</td>
<td>CFP realized in a previous time period.</td>
<td>(1) Accounting-based measures such as ROE in t-1, ROA in t-1, or ROI in t-1 (Galaskiewicz, 1997); (2) market-based measures such as Tobin’s Q in t-1 or market-to-book ratio in t-1 (Brammer et al., 2009a).</td>
</tr>
<tr>
<td>Firm size</td>
<td>The absolute size of the firm.</td>
<td>(1) Absolute measures such as total assets, total sales, total employees, or market capitalization (Al-Tuwaijri et al., 2004).</td>
</tr>
<tr>
<td>Firm leverage</td>
<td>The capital structure of the firm expressed as the balance of debt and equity.</td>
<td>(1) Ratios such as debt/assets, debt/equity, and short term debt/long term debt (Dowell et al., 2000).</td>
</tr>
<tr>
<td>Capital intensity</td>
<td>The balance between capital and labor amongst the firm’s total means of production.</td>
<td>(1) Ratios such as capital expenditure/sales, total assets/sales, number of employees/assets, and sales/assets (King and Lenox, 2002).</td>
</tr>
</tbody>
</table>
relationship in the population of interest’ (Geyskens et al., 2006, p. 526). Compared to Hedges and Olkin meta-analysis (HOMA; Hedges and Olkin, 1985), which only corrects for sampling error, HSMA corrects for a broader array of statistical artefacts, including measurement error, dichotomization, and range restrictions (Heugens et al., 2009), making it a more appropriate choice for estimating mean effect sizes including subjective measures (Kepes et al., 2013). The HSMA method involves (1) estimating the population mean correlation and variability, (2) correcting for the statistical artefacts of sampling and measurement error, and (3) evaluating the existence and impact of moderators (Judge et al., 2011). We performed sub-group analyses to assess whether \( r_c \) was sensitive to: (a) the source of the CSR data (survey, archival, or third-party evaluations); (b) the stakeholder group targeted by the CSR activities (employees, customers, suppliers, shareholders, community, natural environment, or mixed stakeholders); (c) the CFP measures used (accounting- or market-based); and (d) industry (manufacturing, financial services, non-financial services, or mixed industries).

We used both Pearson product-moment correlations (\( r \)) and partial correlations (\( r_{xy.z} \)) derived from primary studies to compute \( \bar{r}_c \) and the corresponding credibility interval. A partial correlation is a standardized measure of the degree of association between two variables (\( x \) and \( y \)), controlling for the influence of a vector of other factors (\( z \)). We obtained the partial correlations from primary studies by converting reported \( t \)-statistics using the relevant formulas.\(^4\) \( r_{xy.z} \)-based HSMA generates useful additional information because they measure the direct impact of CSR on CFP, holding other factors constant (Doucouliagos and Ulubasoglu, 2008). By using \( r_{xy.z} \), we can also compare the mean effect size of the group of primary studies using endogeneity corrections, such as instrumental variables, with that of the group of studies that does not.

When multiple measurements for the same relationship were reported, we included all of them. Monte Carlo simulations have shown that procedures using all reported measurements outperform those that represent each study by a single value in areas such as parameter significance testing and parameter estimation accuracy (Bijmolt and Pieters, 2001). To test whether this nested research design was causing distortions in our results, we conducted a Hierarchical Linear Modeling Meta Analysis (HiLMMA; Raudenbush and Bryk, 2002). The HiLMMA results indicate that 30 per cent of the variance in the effect size distribution resides at the within-study level. After correcting for this component, the mean effect size for the CSR-CFP relationship (i.e., the intercept of the HiLMMA model) is 0.08, which is nearly identical to the uncorrected HSMA mean effect size (0.07; see Table II). Hence, all effect sizes reported in primary studies can safely be included in our analyses.

**MASEM Procedure**

We used meta-analytical structural equation modeling (MASEM; Bergh et al., 2016; Carney et al., 2011) to test our hypotheses. MASEM combines the techniques of structural equation modeling with those of meta-analysis (Cheung and Chan, 2005). This technique allows us to analyse: (a) the direct effect of CSR on CFP, (b) the effect of CSR on firm reputation, stakeholder reciprocation, firm risk, and innovation capacity, and (c) the consequences of these mediating variables for CFP (Baron and Kenny, 1986).
Table II. Hunter-Schmidt meta-analytic results

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Centered Pearson product-moment correlation (r)</th>
<th>Partial linear correlation coefficient (ry.z)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>k</td>
<td>N</td>
</tr>
<tr>
<td>CSR – CFP</td>
<td>1,466</td>
<td>402,863</td>
</tr>
<tr>
<td>Endogeneity control</td>
<td>138</td>
<td>79,979</td>
</tr>
<tr>
<td>Without endogeneity control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSR – CFP without questionnaire</td>
<td>1,280</td>
<td>348,773</td>
</tr>
<tr>
<td>Source of CSR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survey measure</td>
<td>186</td>
<td>54,090</td>
</tr>
<tr>
<td>Archival measures</td>
<td>382</td>
<td>79,555</td>
</tr>
<tr>
<td>Third-party evaluations</td>
<td>898</td>
<td>269,218</td>
</tr>
<tr>
<td>Stakeholder groups</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employees</td>
<td>177</td>
<td>33,531</td>
</tr>
<tr>
<td>Customers</td>
<td>80</td>
<td>17,854</td>
</tr>
<tr>
<td>Suppliers</td>
<td>4</td>
<td>300</td>
</tr>
<tr>
<td>Shareholders</td>
<td>49</td>
<td>10,883</td>
</tr>
<tr>
<td>Community</td>
<td>155</td>
<td>33,098</td>
</tr>
<tr>
<td>Natural environment</td>
<td>412</td>
<td>93,979</td>
</tr>
<tr>
<td>Mixed stakeholders</td>
<td>589</td>
<td>213,218</td>
</tr>
<tr>
<td>Predictor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>k</td>
<td>N</td>
</tr>
<tr>
<td>Measurement of CFP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accounting performance</td>
<td>1,062</td>
<td>270,981</td>
</tr>
<tr>
<td>Market performance</td>
<td>404</td>
<td>131,882</td>
</tr>
<tr>
<td>Industry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manufacturing</td>
<td>177</td>
<td>46,806</td>
</tr>
<tr>
<td>Financial services</td>
<td>89</td>
<td>2,921</td>
</tr>
<tr>
<td>Non-financial services</td>
<td>51</td>
<td>4,833</td>
</tr>
<tr>
<td>Mixed industries</td>
<td>1,149</td>
<td>348,303</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*aSignificant relationships ($p < .05$) are printed in bold.

$b_k = number of samples; N = firm observations; $\bar{r}_c = corrected mean effect size; $\sigma^2_e = sampling error variation; $\sigma^2_p = sample correlation variance; $\hat{\sigma}^2_p = estimated variance in population correlation; CI 95% = 95 percent credibility interval based on $\bar{r}_c$ (see Whitener, 1990; they are partially determined by the amount of estimated variance in population); p-value based on $\chi^2$ tests.

[Correction added on 27 September, 2019, after first online publication: Table II. has been updated.]
MASEM has two advantages over conventional meta-analytic techniques. First, not all specified relationships need to be included in each primary study, as each cell represents an independent accumulation of primary studies (Carney et al., 2011). Second, MASEM allows for the testing of research hypotheses that have never been tested in prior research, especially those connecting previously unlinked streams of literature (Bergh et al., 2016).

MASEM involves a two-stage procedure (Bergh et al., 2016). First, a series of separate r-based HSMA analyses are conducted to construct a meta-analytic correlation table specifying the mean correlations between all independent, mediator, dependent, and control variables. This table is based on bivariate correlations, as partial correlations usually cannot be computed between the independent and control variables included in each primary research study. To deal with sample size differences, we based our analysis on the harmonic mean sample size (N = 3,704). The harmonic mean is less sensitive to outliers than the arithmetic mean, making it more appropriate for estimating correct and conservative t-values (Geyskens et al., 2006).

In the second stage, the meta-analytic correlation matrix is treated as the observed correlation matrix and subjected to structural equation modeling routines (Cheung and Chan, 2005). The data were analysed using the full information maximum likelihood method with the LISREL 8.80 software package. To investigate whether a direct and significant relationship between CSR and CFP exists, and whether this relationship is mediated by our hypothesized variables, we estimated six nested models. We commenced with a model examining the direct relationship between CSR and CFP while controlling for prior firm performance, firm size, firm leverage, and capital intensity. Next, to test our hypotheses using formal tests of statistical mediation (Mackinnon et al., 1995), we added the mediating variables of firm reputation, stakeholder reciprocation, firm risk, and innovation capacity individually (Hypotheses 1 – 4) as well as collectively (Hypothesis 5). This allowed us to trace the change in the coefficient for the direct relationship that resulted from adding the mediators.

RESULTS

Descriptive results: The conceptual structure of Strategic CSR

Following Goertz’ (2006) approach to concept formation, we present in Figure 2 the three-level conceptual structure of Strategic CSR. The concept ‘Strategic CSR’ features at the basic level, at which theoretical propositions are developed such as propositions about the causes and consequences of strategic CSR. At the secondary level, Strategic CSR is intentionally defined in terms of its causally most relevant properties, which in our case are identified by the four empirical mechanisms mediating between CSR and CFP: reputation enhancement, stakeholder reciprocation, risk mitigation, and innovation capacity. At the third, or indicator level of the conceptual structure, we extensionally define Strategic CSR in terms of observable empirical phenomena that operationalize the causally relevant properties by which Strategic CSR is defined. For each causally
relevant property, we list three illustrative indicators (i.e., examples of CSR activities) that in our view are most likely to trigger this mechanism.

The reputation enhancing property of Strategic CSR, for example, captures visible and symbolically-laden CSR activities that may trigger positive attributions from existing and prospective stakeholders. Examples include affiliation with the United Nation’s Global Compact, announcing donations to charities, and voluntary CSR reporting. The stakeholder reciprocation property of Strategic CSR denotes CSR activities that make a substantive contribution to the betterment of the firm’s existing stakeholders. Examples are adopting minority employment policies or family friendly HR practices and signing community benefit agreements. The risk mitigation property of Strategic CSR may involve incremental improvements in practices and processes such as waste reduction and pollution prevention, as well as the adoption of employee health and safety programs. Innovation capacity, finally, captures more comprehensive and transformative CSR initiatives. Potential indicators are public-private partnerships, investments in renewable energy, and base of the pyramid projects.

**Hunter-Schmidt meta-analysis results**

Table II presents the results for our $r$-based (left-hand panel) and $r_{xy.z}$-based (right-hand panel) HSMA analyses pertaining to the CSR-CFP relationship. In addition to the meta-analytic corrected mean effect size ($\overline{r}$), we report the number of samples ($k$), the cumulative sample size ($N$), the sampling error variation ($\sigma_{p}^{2}$), the sample correlation variance ($\sigma_{p}^{2}$), the estimated variance in population correlation ($\hat{\sigma}_{p}^{2}$), the 95 per cent credibility

© 2019 The Authors, Journal of Management Studies published by Society for the Advancement of Management Studies and John Wiley & Sons Ltd.
interval around the $\bar{r}_c$ (CI 95%), as well as the p-value based on $\chi^2$ tests. The $\bar{r}_c$ for the focal relationship is $0.07$ ($k = 1,466$) and the $r_{xy.z}$-based mean is $0.03$ ($k = 1,516$). These means are somewhat smaller than those reported in prior meta-analyses (Margolis et al., 2007:0.13; Orlitzky et al., 2003:0.18). In part, this is due to the fact that our study relies less on self-report data than earlier work (9 per cent vs. 35 per cent for Orlitzky et al., 2003). This has a deflationary effect, since self-report studies usually show larger effect sizes. Yet there is also an independent time effect at play. Even when measured objectively, the effect of CSR on CFP wanes over time ($\bar{r}_c$ is 0.14 for studies published before 1998, but only 0.03 for studies published later), suggesting that the widespread diffusion of CSR practices over the last two decades may undercut some of its value as a strategic differentiating factor.

Sub-group HSMA analyses reveal the presence of methodological artefacts: whereas the $\bar{r}_c$ based on archival and third-party rater data are close to the overall $\bar{r}_c$, those based on survey data are higher ($r$-based mean = 0.13; $r_{xy.z}$-based mean = 0.10). Because few effect sizes are based on survey data (12.7 per cent of the $r$-based effect size distribution and 7.4 per cent of the $r_{xy.z}$-based distribution), removing these effects has no material consequences. When we break down our results by individual stakeholder groups, we find a few effects that are statistically significant but none that seem materially important.

Like Orlitzky and associates (2003), we find CSR to be (somewhat) more strongly linked to accounting-based measures of performance than to market-based measures in the $r$-based distribution ($r$-based mean = 0.07 vs. 0.06). But in the $r_{xy.z}$-based distribution we found CSR to be more linked to market-based measures of performance than to accounting-based measures ($r_{xy.z}$-based mean = 0.04 vs. 0.03). Finally, we found limited evidence for moderation by industry context, with the exception of the financial service industry in the $r$-based distribution ($r$-based mean = 0.23).

**MASEM results and hypothesis tests**

Table III contains the meta-analytic correlation matrix. Each of the 45 cells below the diagonal reports the results of a separate HSMA analysis for each bivariate relationship. We report both the $\bar{r}_c$ and observed correlation ($\bar{r}$). The cells also show the 95 per cent credibility interval, total number of observations ($N$), as well as the total number of samples ($k$) on which the meta-analytic mean is based. Since no primary sample included all correlations included in our model, the total number of samples far exceeds the number of samples in any single cell.

Table IV contains the results pertaining to our Hypotheses 1 through 5. Figure 3 visually presents these same results. The model fits the data well ($\chi^2 = 122.25; \text{RMSR} = 0.024; \text{GFI} = 0.99$). Hypothesis 1 is supported. CSR is positively related to firm reputation ($\beta = 0.10, p < 0.05$), and firm reputation is positively related to CFP ($\beta = 0.05, p < 0.05$). The total indirect effect of CSR that is channelled through firm reputation is 0.005 ($p < 0.05$). A formal test for statistical mediation corroborates these findings (Sobel: $z = 2.65, p < 0.01$). Hypothesis 2 receives marginal empirical support. CSR is positively related to stakeholder reciprocation ($\beta = 0.04, p < 0.05$), which in turn has a positive effect on CFP ($\beta = 0.05, p < 0.05$), but the Sobel test is only borderline significant ($z = 1.77, p < 0.10$) and the total indirect effect of CSR that flows through stakeholder
### Table III. Based Hunter-Schmidt meta-analytic correlation matrix for pooled results

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. CSR</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Prior CFP</td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\bar{r}_c(\bar{r}))</td>
<td>0.05(0.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI 95%</td>
<td>-0.21 / 0.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(N(\bar{r}))</td>
<td>46,804 (132)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Firm size</td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\bar{r}_c(\bar{r}))</td>
<td>0.16 (0.16)</td>
<td>0.13 (0.13)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI 95%</td>
<td>-0.26 / 0.59</td>
<td>-0.23 / 0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(N(\bar{r}))</td>
<td>246,526 (750)</td>
<td>18,799 (52)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Firm leverage</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\bar{r}_c(\bar{r}))</td>
<td>0.01 (0.01)</td>
<td>-0.02 (-0.02)</td>
<td>0.15 (0.15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI 95%</td>
<td>-0.20 / 0.22</td>
<td>-0.38 / 0.33</td>
<td>-0.22 / 0.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(N(\bar{r}))</td>
<td>124,716 (334)</td>
<td>3,019 (19)</td>
<td>75,773 (182)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Capital intensity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\bar{r}_c(\bar{r}))</td>
<td>-0.00 (-0.00)</td>
<td>0.03 (0.03)</td>
<td>0.00 (0.00)</td>
<td>-0.00 (-0.00)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI 95%</td>
<td>-0.28 / 0.28</td>
<td>156 (1)</td>
<td>-0.40 / 0.40</td>
<td>-0.25 / 0.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(N(\bar{r}))</td>
<td>11,534 (54)</td>
<td>5,978 (23)</td>
<td>3,301 (13)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Firm reputation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\bar{r}_c(\bar{r}))</td>
<td>0.14 (0.11)</td>
<td>0.06 (0.05)</td>
<td>0.13 (0.10)</td>
<td>-0.01 (-0.01)</td>
<td>0.03 (0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI 95%</td>
<td>-0.17 / 0.45</td>
<td>-0.10 / 0.22</td>
<td>-0.35 / 0.62</td>
<td>-0.16 / 0.14</td>
<td>-0.08 / 0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(N(\bar{r}))</td>
<td>88,743 (155)</td>
<td>14,281 (9)</td>
<td>37,488 (62)</td>
<td>22,469 (39)</td>
<td>2,372 (9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>7. Stakeholder reciprocation</td>
<td>0.06 (0.05)</td>
<td>0.09 (0.07)</td>
<td>0.10 (0.08)</td>
<td>0.01 (0.01)</td>
<td>0.06 (0.05)</td>
<td>0.06 (0.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ŷ (Ŷ)</td>
<td>0.06 (0.05)</td>
<td>0.09 (0.07)</td>
<td>0.10 (0.08)</td>
<td>0.01 (0.01)</td>
<td>0.06 (0.05)</td>
<td>0.06 (0.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI 95%</td>
<td>−0.30 / 0.41</td>
<td>−0.14 / 0.32</td>
<td>−0.29 / 0.49</td>
<td>−0.11 / 0.14</td>
<td>−0.04 / 0.16</td>
<td>−0.29 / 0.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(k)</td>
<td>45,844 (93)</td>
<td>3,666 (13)</td>
<td>28,158 (55)</td>
<td>26,245 (37)</td>
<td>1,106 (6)</td>
<td>5,303 (13)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Firm risk</td>
<td>0.06 (0.05)</td>
<td>0.06 (0.05)</td>
<td>0.12 (0.10)</td>
<td>0.06 (0.05)</td>
<td>0.12 (0.10)</td>
<td>0.01 (0.01)</td>
<td>0.04 (0.03)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ŷ (Ŷ)</td>
<td>-0.29 / 0.17</td>
<td>-0.21 / 0.33</td>
<td>-0.47 / 0.24</td>
<td>-0.27 / 0.15</td>
<td>-0.22 / -0.01</td>
<td>-0.18 / 0.16</td>
<td>-0.15 / 0.22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI 95%</td>
<td>58,203 (163)</td>
<td>3,562 (18)</td>
<td>24,713 (57)</td>
<td>18,153 (37)</td>
<td>1,171 (5)</td>
<td>13,048 (25)</td>
<td>18,420 (15)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(k)</td>
<td>77,480 (163)</td>
<td>2,423 (14)</td>
<td>41,770 (84)</td>
<td>25,441 (49)</td>
<td>3,161 (12)</td>
<td>20,072 (31)</td>
<td>11,712 (17)</td>
<td>12,987 (23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Innovation capacity</td>
<td>0.08 (0.06)</td>
<td>0.11 (0.09)</td>
<td>0.06 (0.05)</td>
<td>0.14 (0.11)</td>
<td>0.02 (0.02)</td>
<td>0.06 (0.05)</td>
<td>0.21 (0.17)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ŷ (Ŷ)</td>
<td>0.08 (0.06)</td>
<td>0.11 (0.09)</td>
<td>0.06 (0.05)</td>
<td>0.14 (0.11)</td>
<td>0.02 (0.02)</td>
<td>0.06 (0.05)</td>
<td>0.21 (0.17)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI 95%</td>
<td>-0.22 / 0.39</td>
<td>-0.23 / 0.44</td>
<td>-0.43 / 0.31</td>
<td>-0.39 / 0.11</td>
<td>-0.35 / 0.30</td>
<td>-0.18 / 0.29</td>
<td>-0.28 / 0.40</td>
<td>-0.15 / 0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(k)</td>
<td>77,480 (163)</td>
<td>2,423 (14)</td>
<td>41,770 (84)</td>
<td>25,441 (49)</td>
<td>3,161 (12)</td>
<td>20,072 (31)</td>
<td>11,712 (17)</td>
<td>12,987 (23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. CFP</td>
<td>0.07 (0.07)</td>
<td>-0.11 (-0.11)</td>
<td>0.04 (0.04)</td>
<td>-0.06 (-0.06)</td>
<td>0.05 (0.05)</td>
<td>0.07 (0.06)</td>
<td>0.06 (0.05)</td>
<td>-0.11 (-0.09)</td>
<td>0.05 (0.04)</td>
<td></td>
</tr>
<tr>
<td>Ŷ (Ŷ)</td>
<td>0.07 (0.07)</td>
<td>-0.11 (-0.11)</td>
<td>0.04 (0.04)</td>
<td>-0.06 (-0.06)</td>
<td>0.05 (0.05)</td>
<td>0.07 (0.06)</td>
<td>0.06 (0.05)</td>
<td>-0.11 (-0.09)</td>
<td>0.05 (0.04)</td>
<td></td>
</tr>
<tr>
<td>CI 95%</td>
<td>-0.14 / 0.27</td>
<td>-0.67 / 0.44</td>
<td>-0.35 / 0.42</td>
<td>-0.37 / 0.25</td>
<td>-0.18 / 0.28</td>
<td>-0.18 / 0.32</td>
<td>-0.18 / 0.31</td>
<td>-0.38 / 0.17</td>
<td>-0.36 / 0.46</td>
<td></td>
</tr>
<tr>
<td>N(k)</td>
<td>402,863 (1,466)</td>
<td>5,585 (18)</td>
<td>205,145 (569)</td>
<td>138,123 (291)</td>
<td>14,797 (50)</td>
<td>60,806 (101)</td>
<td>74,726 (96)</td>
<td>82,051 (110)</td>
<td>62,432 (129)</td>
<td></td>
</tr>
</tbody>
</table>

*Ŷ = corrected mean effect size; Ŷ = observed correlation; CI 95% = 95 percent credibility interval based on Ŷ; N = firm observations; k = number of samples. Reliability estimates are shown on the diagonal. We used a conservative 0.80 reliability to those variables not verified by independent external auditors who apply legal regulations to attest their accuracy (Bergh et al., 1986).*
Table IV. Hunter-Schmidt (observed correlation) MASEM hypotheses results

<table>
<thead>
<tr>
<th>Predictors</th>
<th>CSR</th>
<th>Firm reputation</th>
<th>Stakeholder reciprocation</th>
<th>Firm risk</th>
<th>Innovation capacity</th>
<th>CFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSR</td>
<td>0.10</td>
<td>0.04</td>
<td>-0.04</td>
<td>0.07</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Prior CFP</td>
<td>0.03</td>
<td>0.06</td>
<td>-0.04</td>
<td>0.07</td>
<td>-0.13</td>
<td></td>
</tr>
<tr>
<td>Firm size</td>
<td>0.16</td>
<td>0.07</td>
<td>-0.10</td>
<td>-0.06</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Firm leverage</td>
<td>-0.02</td>
<td>0.00</td>
<td>-0.03</td>
<td>-0.10</td>
<td>-0.07</td>
<td></td>
</tr>
<tr>
<td>Capital intensity</td>
<td>0.02</td>
<td>0.05</td>
<td>-0.10</td>
<td>-0.02</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Firm reputation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Stakeholder reciprocation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Firm risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.09</td>
<td></td>
</tr>
<tr>
<td>Innovation capacity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
<td></td>
</tr>
</tbody>
</table>

- $R^2$ | 0.026 | 0.021 | 0.014 | 0.027 | 0.026 | 0.042 |
- Harmonic mean $N$ | 3,704 |
- $X^2$ | 122.25 |
- $GFI$ | 0.99 |
- $RMSR$ | 0.024 |

*Significant relationships ($p < 0.05$) are printed in bold; $t$-values are given in parentheses.
reciprocation amounts to only 0.002 (p < 0.05). Hypothesis 3 is supported. CSR is negatively related to firm risk (β = −0.04, p < 0.05), and firm risk is negatively related to CFP (β = −0.09, p < 0.05). The total indirect effect of CSR that is channelled through firm risk is 0.004 (p < 0.05) and the Sobel test is significant (z = 2.08, p < 0.05). We also found support for Hypothesis 4. CSR has a positive and significant effect on innovation capacity (β = 0.07, p < 0.05), and innovation capacity positively affects CFP (β = 0.05, p < 0.05). The total indirect effect of CSR that flows through innovation capacity is 0.004 (p < 0.05) and the Sobel test is significant (z = 2.53, p < 0.05). Hypothesis 5 is rejected, however. On the one hand, the Sobel test confirms that the relationship between CSR and CFP is significantly channelled through the four hypothesized pathways (z = 4.51, p < 0.001), with the total indirect effect amounting to 0.01 (p < 0.01). On the other hand, firm reputation, stakeholder reciprocation, firm risk, and innovation capacity only partially mediate the relationship between CSR and CFP, as the focal effect remains significant (β = 0.06, p < 0.05) upon their inclusion. Specifically, inspection of the path coefficients for the indirect effects suggests that the four hypothesized mediators combined explain 20.0 per cent of CSR-CFP relationship. This suggests that there may be alternative causal paths connecting CSR to CFP.

**Additional analyses**

We performed several additional analyses to address a number of problems that commonly affect meta-analyses. The first of these is the ‘file drawer’ problem, or the under-representation of studies reporting weak effects due to publication bias. To address it, we crafted r- and r_{xy,z}- based funnel plots of effect and sample sizes (see Figure 4a and b). These figures show a symmetrical pattern with nearly half of the samples to the left of...
the mean effect size and half to the right; thus suggesting the absence of publication bias (Duval, 2005). We also used the triangulation method to detect the possible presence of publication bias in the CSR-CFP relationship (Harrison et al., 2017). This method involves three complementary tests to detect potential differences between the retrieved HSMA mean effect size and a publication bias-adjusted mean effect size estimate: (a) Duval and Tweedie’s trim and fill (Duval, 2005), (b) cumulative meta-analysis (CMA), and (c) moderate selection models (Harrison et al., 2017). The results of these tests suggest the absence of publication bias in our meta-analysis.

A second common problem is undetected measurement error in the underlying primary studies. We already applied a 0.80 level of reliability to all our mediating variables in the main results (see Tables III and IV). As an additional robustness check, we re-ran our models using the 0.80 level of reliability for all constructs (Bergh et al., 2016). The findings of this more conservative analysis are consistent with those presented in Table IV.

A third common problem in meta-analyses of literatures based on archival data is overlapping primary samples. We therefore ran a robustness check in which we deleted effect sizes stemming from similar sources (e.g., Fortune 500 firms) with comparable median sampling years. This resulted in the removal of 39 $r$-based studies and 16 $r_{xy,z}$-based studies. Our results were not materially affected by the removal of these overlapping samples.

A fourth issue hampering meta-analyses that also partially include survey data is that self-report studies tend to report larger effect sizes, possibly due to social desirability and self-aggrandizing biases. We therefore ran a separate MASEM analysis on effects derived from studies reporting objective data. Our results were identical to those of the original analysis.

A fifth recurring problem affecting especially MASEM analyses is that they tend to be based largely on cross-sectional studies lacking appropriately lagged variables, which might violate the implicit assumption of temporal sequentially underlying path analysis. In our sample, 40 per cent of all studies are in fact based on longitudinal designs. As an additional robustness check, we therefore re-ran our hypothesized model on the subset of effect sizes derived from this set of studies. The results rejected the mediation effect of firm reputation (Hypothesis 1), possibly due to statistical power attrition, but confirmed Hypothesis 2, 3 and 4.

A sixth issue, again specific to MASEM studies, is that the theorized model may not be the model that empirically fits the data best. To account for this possibility, we tested two alternatively specified models (see Bergh et al., 2016). In a first alternative model, we related (a) CSR to reputation enhancement and stakeholder reciprocation; (b) reputation enhancement to CFP; and (c) stakeholder reciprocation to CFP via firm risk and innovation capacity. In a second alternative model, we related (a) CSR to firm reputation, stakeholder reciprocation, firm risk, and innovation capacity; (b) firm reputation to firm risk and innovation capacity; and (c) stakeholder reciprocation to firm risk and innovation capacity, and (d) firm reputation, stakeholder reciprocation, firm risk, and innovation capacity to CFP. We found the theorized model to show better performance on 10 out of 14 commonly used model fit indicators.
In sum, we conclude that our results are largely robust against a number of issues that occasionally affect the reliability and validity of meta-analytic findings. The results of all robustness tests are available in an online appendix7.

DISCUSSION

After reviewing almost half a century of empirical research on the CSR-CFP relationship, we have developed hypotheses on the causally most relevant empirical mechanisms through which CSR has been found to positively affect CFP. We subsequently used MASEM to test these hypotheses on the accumulated empirical evidence. As such, we aim to make several contributions to the literature.

First, our study offers a new synthesis of the available evidence that includes the last two decades of CSR-CFP research. In the fifteen years since the publication of the last two meta-analyses on the relationship between CSR and CFP (Margolis and Walsh, 2003; Orlitzky et al., 2003), the field of empirical CSR research has grown significantly as evidenced by the fact that our analysis counts no less than 224 new papers. During this period, the focus has shifted from empirically assessing the CSR-CFP relationship...
as a whole to examining the more concrete empirical mechanisms underlying it. Given this development, an up to date assessment of the accumulated empirical evidence is timely and useful. Next to identifying and testing four key mechanisms through which CSR activities affect CFP, our meta-analysis also found that these mechanisms do not fully mediate the CSR-CFP relationship. This implies that there is still room for future research to identify empirical mechanisms causally connecting CSR with CFP that have gone unobserved to date. We return to this below.

A second contribution of our study is that we respond to calls for more conceptual clarification of CSR (Gond and Crane, 2010; Rowley and Berman, 2000; Van Oosterhout and Heugens, 2008). By focusing on the empirical relationship that has received most scholarly attention to date (Margolis and Walsh, 2003), we were able to distinguish and conceptually carve out Strategic CSR from CSR more generally, and provide a definition of Strategic CSR that is both intentionally and extensionally explicit (Van Oosterhout, 2005). Specifically, we define Strategic CSR as those firm activities that appear to further some social good, while at the same time benefiting the firm financially by either enhancing its reputation, increasing stakeholder reciprocation, mitigating firm-specific risk, and/or improving innovation.

With this evidence-based conceptualization of Strategic CSR, we address an important conceptual problem that has burdened much of the CSR literature to date. By positioning a firm’s economic responsibilities as one of a larger set of social responsibilities that comprise CSR (Carroll, 1979), a central critique of the CSR concept has always been that it confounds two fundamentally conflicting views: that of positive economics and normative ethics (Gond and Crane, 2010; Jones, 1983; Swanson, 1995; Windsor, 2006). Some scholars have indeed argued that CSR is a normative concept (Matten et al., 2003) that has little positive explanatory value (Van Oosterhout and Heugens, 2008). In this study, we have conceptually developed the Strategic CSR concept that can be used for positive research in strategic management, however. This concept only bears on CSR activities that enhance CFP, making Strategic CSR a purely instrumental form of CSR. By excluding those CSR activities that may be good for society but not for the firm, we have sought to conceptually separate some of CSR’s main positive explanatory features from the normative beliefs as to what CSR ought to comprise. By disentangling the normative and positive features of CSR in this way, we hope to provide some relief to a field that has consistently confused normative and positive theorizing, and which has been a major impediment to its development (Gond and Crane, 2010).

The development of a conceptualization of Strategic CSR does not mean that the discussion on what Strategic CSR is, is now completed, nor that all issues on what CSR should be taken to mean more generally have now been resolved. With respect to the former, it should be clear that future empirical research may still lead to further evidence-based adjustments of the Strategic CSR concept, as would be true in general for all causal approaches to explanatory concept formation. With respect to the latter, our study has conceptually clarified Strategic CSR by carving it out of the domain denoted by CSR more generally, but only at the price of leaving that domain as conceptually opaque as it was (claimed to be) before. Both observations may help to guide future research on CSR.
Limitations and future research

A third contribution of this study, therefore, is that its findings and limitations offer guidance to future research. A salient finding of our study is that the four mechanisms combined explain only 20 per cent of the CSR-CFP relationship. These mechanisms therefore do not fully mediate the relationship between CSR and CFP, suggesting that our conceptualization of Strategic CSR does not yet fully account for the mechanisms connecting CSR to CFP. One explanation for this may be that research to date has not yet captured the full breadth and depth of each of the four mechanisms identified. With respect to stakeholder reciprocation, for example, we found studies on employee, government, customer, and shareholder reciprocation, but hardly any studies on community reciprocation (but see: Dorobantu and Odziemkowska, 2017), and none on supplier reciprocation. Similarly, the benefits of stakeholder reciprocation may become manifest in different guises, including retention, word of mouth, and commitment, which all require further empirical measurement and investigation (Bosse and Coughlan, 2016). As more empirical research fleshing out these four mechanisms accumulates, however, we expect the explanatory power of each of these mechanisms to strengthen further.

An alternative explanation for our partial mediation findings may be that existing research has not yet unveiled all empirical mechanisms through which CSR may contribute to CFP. There may be other mechanisms at play than the four that we identified in this study, implying that our definition of Strategic CSR may need to be expanded and adjusted. As observed above, this is consistent with a causal approach to explanatory concepts in social science, in which new empirical knowledge about a phenomenon may require adjustments or expansions in its conceptual definition (Goertz, 2006).

Using meta-analysis to investigate our research question is one of the strengths of our paper, but it also has limitations. While our conceptualization of Strategic CSR assumes that the empirical mechanisms that we document involve causal relationships, meta-analysis is not the best method for identifying causality in empirical research. Rather, meta-analysis is more suited for investigating which theoretical views about causal relationships are supported by a preponderance of the available evidence. The growing and increasingly fragmentated nature of existing CSR-CFP research makes such an assessment of the literature imperative. By identifying the four causally most relevant empirical mechanisms underlying the CSR-CFP relationship, our study may therefore constitute a counter-force to further fragmentation of the field, and provide guidance to researchers seeking to address ever finer-grained research questions in order to identify causality through their research designs.

A related limitation of meta-analysis is that the effect sizes that can be included as data depend on their availability in primary studies. Because available empirical research on the CSR-CFP relationship is highly diverse, we were able to incorporate many effect sizes and empirically test finer-grained models than have been meta-analytically tested to date. Yet we could not address any questions for which effect sizes were not or insufficiently available in primary studies. This problem is most pressing with respect to the measurement of CSR itself, which prior research has often operationalized by using overly inclusive indicators that lump together many different phenomena (Chatterji et al., 2016; Gond and Crane, 2010; Rowley and Berman, 2000; Van Oosterhout and
Heugens, 2008). If anything, we urge researchers to use more specific measurements of Strategic CSR, derived from the empirical mechanisms that are or have been investigated (or, to say the same in other words: to adopt level 3 indicators that derive from secondary level mechanisms, as illustrated in Figure 2).

Finally, we acknowledge that our study offers conceptual clarification on the part of the CSR concept that has strategic value for the firm. As a result, our study says very little about CSR that does not benefit the firm financially. Some scholars have explicitly argued that managers should *only* engage in CSR if it simultaneously advances the bottom line (Jensen, 2002; Waldman and Siegel, 2008). We do not agree, however, that all CSR is or ought to be, strategic. Some CSR activities may have a neutral or insignificant impact on CFP, while others may have a significant negative effect, at least in the short to medium term (Wright and Ferris, 1997). Our meta-analysis did not touch upon the non-strategic side of the CSR concept as this has mostly been a blind spot to CSR scholars. To understand and conceptualize non-strategic CSR, empirical research is needed that explores when and why managers sometimes pursue CSR activities that are not directly financially beneficial to the firm. Do such activities only serve managerial self-interests (Jensen, 2002; Waldman and Siegel, 2008), or are managers driven by more altruistic motives (Husted and Salazar, 2006)? Moreover, under what conditions do a firm’s stakeholders, especially its shareholders, accept non-strategic forms of CSR? This will likely depend on the consequences of non-strategic CSR, not only for the firm, but also for society. The literature currently defines CSR as all firm activities that *appear* to further some social good (McWilliams and Siegel, 2001), but CSR scholarship has mostly ignored the *empirical* question of whether CSR *actually* creates value for its intended beneficiaries or not (Margolis and Walsh, 2003). As the value of CSR should not be evaluated on the basis of its strategic benefits for the firm alone, we hope that the next wave of CSR research will focus more on empirically investigating the field-defining assumption that CSR furthers some form of social good and thereby increase our understanding of non-strategic forms of CSR.

**NOTES**

[1] Overall, the causal approach to concept formation that we subscribe to therefore relies on a *functionalist* rather than *essentialist* understanding of concepts. In this functionalist view, concepts are understood in terms of what they do (e.g., how they explain reality) instead of what is assumed about the nature of reality that the concept refers to (e.g., what that part of reality is essentially like even if there exists no feasible way of ever truly knowing what it is like). In this view, which Goertz (2005) develops for social science concepts, our conceptual understanding of a phenomenon therefore evolves with our empirical understanding of how the phenomenon denoted by the concept interacts with other phenomena.

[2] One might object, as an anonymous reviewer did, that ‘hardwiring’ causal knowledge on the consequences of CSR into the conceptualization of strategic CSR would frustrate these causal claims from empirically being tested. This is a legitimate concern for empirical research indeed, as has previously been explained in the literature (Mackenzie, 2003; Mackenzie et al., 2011) and which has also been acknowledged by Goertz (2009, p. 65–67). For three reasons, however, we do not believe this concern to burden our evidence-based conceptualization of Strategic CSR, however. First, as we have indicated in the body of the text, developing strategic concepts by inferring them from performance differentials between firms is well-established in the field of strategic management, given its field-defining objective of explaining firm competitive advantages from observed performance.
differentials between firms. In any academic field, however, an important part of conceptually understanding a phenomenon will result from making sense of how the phenomenon causally influences other phenomena. Second, as the main aim of this study is to develop the concept of strategic CSR based on the body of causal knowledge and empirical evidence already available at this time, we are not overly worried about these claims being precluded from empirical testing. Not only do we provide a comprehensive meta-analytic test of all the mechanisms that we use to identify the causally most relevant properties of CSR in explaining CFP, but all causal claims involved in our conceptualization of Strategic CSR are open to later empirical scrutiny should there be substantive or methodological reasons for empirically testing these claims further. Finally, Goertz (2009, p. 65–7) acknowledges that hardwiring causal knowledge into our theoretical conceptualization of social reality may not only increase our theoretical understanding of social reality, but may also be problematic precisely because conceptual knowledge is assumed rather than tested in any particular empirical research project. He proposes a pragmatic approach to this problem, however, by arguing that hardwiring available causal knowledge into our conceptualizations should be avoided as much as possible when conceptualizing the explanandum (or dependent variable) in research. This is because in explaining a phenomenon, one should strive to make all explanations of this phenomenon as explicit as possible, in order to be able to empirically assess and test all of them separately and against each other. These concerns are mitigated, however, when conceptualizing the explanans (or explanatory variable) in research, because in that case we would be wise to use all currently available causal knowledge in our concepts and theories in order to develop and test novel theories and predictions. Note that our evidence-based conceptualization of Strategic CSR treats it as an explanans in explaining CFP, while we do not make any explanatory claim in this study about Strategic CSR as an explanandum. In sum, we believe our strategy to conceptualize Strategic CSR based on available knowledge of how it causally affects CFP not only increases our theoretical understanding of Strategic CSR, but also our understanding of how Strategic CSR differs from CSR more generally.

[3] A complete overview of the included studies in our meta-analysis can be offered upon request.

[4] The partial correlation coefficients were calculated by using the t-statistic reported in the primary studies. If the t-statistic was not reported, we approximated it using the regression coefficient and the standard errors. The formula used to calculate partial correlation is: \( r_{xy.z} = \frac{t}{\sqrt{(t^2 + df)}} \), where t is the t-statistic and df is degrees of freedom. Note that this will always produce a positive number, so it is necessary to convert it to a negative number if the regression coefficient is negative (Greene, 2008).

[5] As a separate robustness check, we also ran our MASEM analyses again without the survey data-based effect sizes. All our findings are robust against this exclusion.

[6] For the Duval and Tweedie’s trim and fill method, we used the left-of-mean (L) estimator (Kepes et al., 2012). We then calculated the absolute change in expected direction from the Hunter and Schmidt mean effect size to the left of the trim and fill-adjusted observed mean effect size (Harrison et al., 2010). For the CMA, we calculated the weighted mean correlation of the 10 per cent most precise (i.e., largest samples) in our meta-analysis (Borenstein et al., 2009; Harrison et al., 2010). Again, we calculated the absolute change in expected direction from the Hunter and Schmidt mean effect size, but now to the most precise 10 per cent of samples (Harrison et al., 2010). For the moderate selection model, we followed the Hedges and Vevea (2005) technique, since several of our meta-analytic relationships present a k lower than 100. We identified the specific weights and probability value cut-off points for our sample based on Vevea and Woods (2005). The absolute change in expected direction from the Hunter and Schmidt mean effect size was calculated to the moderate selection model adjusted mean effect size (Harrison et al., 2010). We included the absolute change in expected direction from the Hunter and Schmidt mean effect size to: (a) the trim and fill adjusted observed mean effect size, (b) the most precise 10 per cent of samples (CMA), and (c) the moderate selection model. We then calculated the average degree of change across these three publication bias tests (Harrison et al., 2010), which was equal to 0.01. We also computed Rosenthal’s Fail-Safe N values: \( N = 829,029 \) and 436,029 for the r- and \( r_{xy.z} \)-based correlation results respectively. These numbers represent the number of missing studies averaging a z-value of zero that should be added to our sample to make the combined effect size statistically insignificant.

[7] The online appendix is available at: https://figshare.com/s/088ee166da7c37693b0b
REFERENCES


