European Management Journal 32 (2014) 735–751

Contents lists available at ScienceDirect

European Management Journal

journal homepage: www.elsevier.com/locate/emj

Making sense of conflicting empirical findings: A meta-analytic review of the relationship between corporate environmental and financial performance

Jan Endrikat [a,1], Edeltraud Guenther [b,*], Holger Hoppe [c,2]

[a] Technische Universitaet Dresden, Faculty of Business and Economics, Chair of Business Management, esp. Management Accounting and Control, 01062 Dresden, Germany
[b] Technische Universitaet Dresden, Faculty of Business and Economics, Chair of Business Management, esp. Environmental Management and Accounting, 01062 Dresden, Germany
[c] Linde Material Handling GmbH, Carl-von-Linde Platz, 63743 Aschaffenburg, Germany

ARTICLE INFO

Article history:
Received 1 July 2013
Accepted 17 December 2013
Available online 13 January 2014
Handling Editor: Steven Grover

Keywords:
Corporate environmental performance
Corporate financial performance
Meta-analysis
Causality
Multidimensionality
Moderation effects

ABSTRACT

Despite the tremendous number of publications concerned with the relationship between corporate environmental performance (CEP) and corporate financial performance (CFP), inconsistent empirical findings persist and the overall picture remains vague. Drawing on a hybrid theoretical framework (combining the theoretical reasoning of the natural-resource-based view (NRBV) with instrumental stakeholder and slack resources arguments), we address the apparent lack of consensus by meta-analytically integrating the findings of 149 studies. We pay particular attention to two highly material issues: the direction of causality and the multidimensionality of the focal constructs. Meta-analytic results indicate that there is a positive and partially bidirectional relationship between CEP and CFP. In addition, our findings suggest that the relationship is stronger when the strategic approach underlying CEP is proactive rather than reactive. Furthermore, we reveal moderation effects of methodological artifacts, which may provide explanations for the inconsistency of the results of previous studies. Based on our findings, we discuss the implications and outline avenues for further research.

Introduction

> "While the question of whether it pays to be green has probably generated more research pages than any other single question, the answer remains unresolved" (Hoffman & Bansal, 2012, p. 14).
>
> "Science advances when scholars reach consensus about the conclusions offered by a body of evidence, and meta-analysis is our best methodology for reaching consensus" (Combs, Ketchen, Crook, & Roth, 2011, p. 194).

The relationship between corporate environmental performance (CEP) and corporate financial performance (CFP) constitutes one of the most puzzling phenomena pertaining to research on organizations and the natural environment. Insofar as scholars have tackled environmental issues, the discovery of the link between CEP and CFP has evolved into something similar to finding the "holy grail" (Boons & Wagner, 2009; Peloza, 2009). Over the last four decades, myriad studies have sought to identify the relationship between these performance constructs. In this context, one of the most fundamental issues shaping research on the focal relationship refers to the direction of causality (i.e., whether CEP influences CFP, whether CFP influences CEP, or whether there is a bidirectional relationship) (e.g., Ambec & Lanoie, 2008; Molina-Azorín, Claver-Cortés, López-Gamero, & Tarí, 2009; Preston & O'Bannon, 1997). Furthermore, in recent years, several scholars (e.g., Aragón-Correa & Sharma, 2003; Delmas, Hoffmann, & Kuss, 2011; Etzion, 2007; Orlitzky, Siegel, & Waldman, 2011) have emphasized the need to extend the scope of analysis by exploring the determinants, potential contingency factors, and boundary conditions under which CEP and CFP are related. In other words, research has been called upon to adopt a more sophisticated view and to "look beneath the surface" (Delmas et al., 2011, p. 117) in order to get a better understanding of the mechanisms connecting both performance constructs.

Despite the enormous number of publications concerned with the relationship between CEP and CFP, the overall picture remains vague. While some studies have provided evidence of a positive relationship (e.g., Clarkson, Li, Richardson, & Vasvari, 2011; Hart & Ahuja, 1996; King & Lenox, 2001; Konar & Cohen, 2001; Russo & Fouts, 1997; Wagner & Schaltegger, 2004), others have sup-

* Corresponding author. Tel.: +49 351 463 34313; fax: +49 351 463 37764.
E-mail addresses: jan.endrikat@tu-dresden.de (J. Endrikat), ema@mailbox.tu-dresden.de (E. Guenther), holger.hoppe@linde-mh.de (H. Hoppe).
[1] Tel.: +49 351 463 37015; fax: +49 351 463 37712.
[2] Tel.: +49 6021 991480; fax: +49 6021 996480.

0263-2373/$ - see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.emj.2013.12.004

ported the conclusion of a negative relationship (e.g., Cordeiro & Sarkis, 1997; Hassel, Nilsson, & Nyquist, 2005; Morris, 1997) or yielded insignificant results (e.g., Cohen, Fenn, & Konar, 1997; Graves & Waddock, 1999). Several explanations for the apparent inconsistency have been proposed, involving both methodological and theoretical issues (Ruf, Muralidhar, Brown, Janney, & Paul, 2001). These explanations address different aspects, describing (1) the lack of a sound theoretical foundation (e.g., Aragón-Correa & Sharma, 2003; Ullmann, 1985); (2) the lack of a clear idea of the direction of causality (e.g., Ambec & Lanoie, 2008; Surroca, Tribó, & Waddock, 2010; Waddock & Graves, 1997); (3) the inconsistency of defining and measuring the constructs of interest (e.g., Busch & Hoffmann, 2011; Griffin & Mahon, 1997; McWilliams, Siegel, & Wright, 2006); and (4) the use of misspecified models due to omitted variables and a lack of consideration of moderating or mediating influences (e.g., Russo & Minto, 2012; Telle, 2006).

Prior reviews and meta-analyses: Rationale for the present study

Inconclusive findings in a given field of inquiry (as it is the case for the CEP-CFP nexus) constitute great nuisances to researchers by limiting the understanding of certain phenomena and impeding credible scientific generalizations (e.g., Rousseau, 2006). Thus, it is not surprising that several attempts have been made to consolidate the empirical research on the relationship between CEP and CFP. In order to clarify the distinct contribution of our study, we briefly review these prior works and explain why they merely allow limited conclusions regarding the focal relationship.

Narrative reviews, vote counts and the superiority of meta-analysis

The vast majority of existing reviews on the relationship between CEP and CFP either used narrative approaches or applied vote count procedures. Narrative reviews, provided for example by Ambec and Lanoie (2008), Blanco, Rey-Maquieira, and Lozano (2009), Guenther and Hoppe (2014), Molina-Azorín et al. (2009), and Salzmann, Ionescu-Somers, and Steger (2005), summarize the available research in a descriptive manner without providing a quantitative integration. These reviews without any doubt have contributed to the integration of the vast amount of CEP–CFP studies. However, without calling into question the general appropriateness of narrative reviews, at least in cases where multiple studies have yielded inconclusive results, narrative approaches are subject to several limitations (e.g., purely descriptive nature, subjectivity, and lacking critical assessment) and thus hardly enable conclusions that make sense of conflicting empirical findings (e.g., Hart, 1998; Tranfield, Denyer, & Smart, 2003). Another approach that has been applied for synthesizing the body of empirical CEP–CFP studies refers to the vote count technique (e.g., Guenther, Hoppe, & Endrikat, 2011; Horváthová, 2010; Margolis & Walsh, 2001).[3] In vote counts study findings are simply coded and aggregated as positive, negative, or non-significant (Orlitzky, Schmidt, & Rynes, 2003). This approach has been strongly criticized by many management scholars and statistical experts due to several substantial problems (e.g., Type II error problems, ignoring of sample size differences, or lacking provision of a point estimate of effect sizes), which likely lead to invalid conclusions (Combs et al. 2011; Dalton & Dalton 2005; Orlitzky et al. 2003).[4] In contrast to narrative reviews or vote counts meta-analytic methodology, which is based on accurate statistical aggregation, is the most sophisticated research-integration technique providing a sound way to quantitatively accumulate empirical findings, to make sense of inconclusive empirical evidence, and to draw conclusions that reconcile conflicting results (e.g., Aguinis, Dalton, Bosco, Pierce, & Dalton, 2011; Combs et al., 2011; Geyskens, Krishnan, Steenkamp, & Cunha, 2009; Orlitzky et al. 2003).

Previous meta-analyses

Orlitzky et al.'s (2003) prominent meta-analysis explored the more general relationship between corporate social performance (CSP) and CFP.[5] Their study sample covered the years from 1972 to 1997 and only included US studies. By breaking down the entire sample of studies in different subgroups they examined, among other points, the relationship between CEP and CFP (139 effect sizes). Allouche and Laroche (2005) provided another meta-analysis including 82 studies concerned with the relationship between CSP and CFP. They also partly captured the CEP–CFP link by means of subgroup analysis (84 effect sizes).

There are several reasons why the existing CSP–CFP meta-analyses hardly allow drawing definitive conclusions with regard to the relationship between CEP and CFP. First, the number of CEP–CFP studies included in these analyses is far from exhaustive. That applies in particular to the influential study of Orlitzky et al. (2003), which included at the most only 17 CEP–CFP studies.[6] Furthermore, their analysis only included studies published until 1997. Given that knowledge about social phenomena, in general, tends to be historically dependent (Combs et al., 2011; Gergen, 1976), and that environmental issues in particular have gained increasing importance during the last years, it is likely that study findings vary over time and that more recent studies that have not been included yielded different results. Second, despite the lack of a commonly shared definition or conceptualization of CEP (e.g., Etzion, 2007; Walls, Phan, & Berrone, 2011; Xie & Hayase, 2007), it is beyond question that CEP is a multidimensional construct (e.g., Clemens & Bakstran, 2010; Dooley & Fryxell, 1999; First & Khetriwal, 2010; Trumpp, Endrikat, Zopf, & Guenther, in press), which has been measured in several different ways. However, none of the CSP–CFP meta-analyses decomposed the CEP construct into more specific dimensions and thus failed to account for the multidimensional nature of CEP and possible moderation effects concerning this matter. Third, in all CSP–CFP meta-analyses the moderator analyses as well as analyses with regard to the direction of causality have been conducted only at the level of the overall analysis pertaining to CSP. As CEP constitutes one dimension of the broader (meta-) construct of CSP (Orlitzky et al., 2003), there are undoubtedly similarities regarding theoretical reasoning and empirical manifestations. However, methodological literature on multidimensional constructs urgently cautions that confounding construct level (in this case CSP) and dimension level (in this case CEP) would imply serious threats to validity (e.g., Law, Wong, & Mobley, 1998; Wong, Law, & Huang, 2008). Indeed, notwithstanding the similarities, there are also fundamental differences that may warrant a separate treatment of CEP and CSP (Bansal & Gao, 2006). Walls, Berrone, and Phan (2012, p. 892), for example, argued that environmental issues tend to differ from social issues because "they are technical, require specific firm capabilities and significant capital investment, are guided by regulation, and have their own reporting criteria". Also empirical examinations have shown that CEP should be considered a distinct part of the overarching construct of CSP and that certain levels of CSP do not necessarily correspond to sim-

[3] While Horváthová (2010) labeled her study as a meta-analysis, she obviously applied a vote count procedure.

[4] For a detailed discussion pertaining this issue see for example Combs et al. (2011) or Hedges and Olkin (1985).

[5] As CEP constitutes a subdimension of the more general concept of CSP (Orlitzky et al., 2003), there is a strong intermingling of research on CEP and research on CSP. Throughout this paper, when analogies are obvious and appropriate, we draw on arguments and findings that refer to the broader construct of CSP.

[6] The exact number cannot be determined due to missing information.