International Journal of Production Economics 219 (2020) 328–340

Contents lists available at ScienceDirect

International Journal of Production Economics

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

Unraveling the complex relationship between environmental and financial performance —— A multilevel longitudinal analysis

Zuoming Liu

3820 Mundy Mill Rd, Mike Cottrell College of Business, University of North Georgia, Oakwood, GA 30566, USA

ARTICLE INFO

Keywords:
Environmental performance
Financial performance
Green operations
Sustainability
Multilevel

ABSTRACT

Many business managers start to adopt environmental-friendly activities due to pressures and concerns regarding the potential adverse environmental impacts from their regular activities. However, there is no consensus regarding whether better environmental performance can lead to superior financial performance. Academic research also shows mixed or even contradictory results, possibly due to various limitations and problems. This study builds a multilevel framework to study the complicated relationship between environmental performance and financial return and synthesize those non-consensus results in many previous studies. An overall positive relationship is found between environmental and financial performance with variations across companies and industries due to the company- and industry-level heterogeneities. A negative relationship is also identified for some firms in some industry sectors. Moreover, the bi-directional causal relationship between environmental and financial performance also implies that it is necessary for companies to have sufficient financial resources in order to implement proactive environmental strategies and initiatives. The findings in this study make contributions to the literature, provide guidelines for managers and investors, and give implications for policymakers.

1. Background

As society has been paying more and more attention to the negative impacts of many business operations on the environment, firms are facing increasing pressures to improve their social responsibility and achieve sustainable development. As an important component of corporate social responsibility (CSR), environmental management (EM) has been gaining soaring attention from consumers, businesses, governments, non–governmental organizations (NGO) and academics as well. Many stakeholders have been starting to urge firms to reduce their existing or potential negative impacts on the natural environment, community, and society. Numerous regulations and policies have been proposed by government agencies and NGOs aiming to reduce or prevent environmental deterioration. Accordingly, an increasing number of firms started to adopt EM practices under compulsion or voluntarily. For example, based on the study from KPMG, about two-thirds of the largest companies in western countries have engaged in green or sustainable development to some extent and published related environmental disclosure reports (KPMG, 2011). The costs related to EM activities have kept increasing substantially (Barbera and McConnell, 1990).

How does a firm's EM affect its financial performance (FP)? Does it pay to be green? After decades of theoretical and empirical studies, there still seems no conclusive results (Konar and Cohen, 2001; Wagner et al., 2001). Overall, two opposite opinions exist. On the one hand, the neoclassical school of economics argues that additional costs will be incurred for firms from conforming environmental regulations (Walley and Whitehead, 1994). Thus, expenditures and efforts devoted to EM practices to improve environmental performance (EP) are usually viewed as extra costs that will decrease financial returns. On the other hand, drawing on resource-based view (RBV), Porter's hypothesis and stakeholder theory, researchers suggest a positive relationship between EP and FP because of increased legitimacy and sustainability (Clarkson, 1995; Porter, 1991; Hart, 1995; Porter and van der Linde, 1995; Li et al., 2019). In strategic management studies, RBV has been used for decades to explain corporation competitive advantages by arguing that firm's rare and valuable resources or capabilities are hard to imitate or substitute (Wernerfelt, 1984). Hart (1995) posits that "capabilities that facilitate environmentally sustainable economic activity" (p. 991) is critical for a firm when building competitive advantages with the acceleration of negative impacts on natural resources and environment in both scale and scope from human activities. Porter's hypothesis and stakeholder theory argue that extra benefits such as cost reduction and innovations from EM can offset the costs of complying with environmental regulations with improved efficiency and enhanced competitive advantage.

As a result, mixed outcomes have been found in the literature regarding the relationship between EP and FP (Horváthová, 2010;

E-mail address: Zuoming.Liu@ung.edu.

https://doi.org/10.1016/j.ijpe.2019.07.005
Received 30 October 2018; Received in revised form 3 July 2019; Accepted 9 July 2019
Available online 10 July 2019
0925-5273/

Orlitzky et al., 2003). First of all, some studies found that EM contributes significantly to FP, i.e., it pays to be green (Christmann, 2000; Judge and Douglas, 1998; King and Lenox, 2001; Klassen and Whybark, 1999; Konar and Cohen, 2001; Russo and Fouts, 1997). Second, although some other studies demonstrated the significant association between EP and FP, it only exists when firms were not doing well in their EP with negative environmental issues or accidents (Beatty and Shimshack, 2010; Gupta and Goldar, 2005). Third, there are also some studies showed no significant association between EP and FP (Aras et al., 2010; Moneva and Cuellar, 2009; Waddock and Graves, 1997). Lastly, many others indicated a negative relationship between EP and FP (Jacobs et al., 2010; Konar and Cohen, 2001; Moore, 2001; Ngwakwe, 2009; Stanwick and Stanwick, 1998).

Various explanations have been proposed trying to explain these inconsistent findings, such as small sample size (Konar and Cohen, 2001), lack of objective criteria to evaluate EP as well as lack of clear definitions for EP and FP (Cohen et al., 1995; Griffin and Mahon, 1997), self–reported cross-sectional data with large common method of variance (Wood and Jones, 1995), omissions of important contingent factors such as firm size or location (Elsayed and Paton, 2005; Wagner et al., 2001), negligence of firm and industry heterogeneities and over–time dynamics (Elsayed and Paton, 2005), and methodology limitations failing to address the complicated structure between EP and FP (Derwall et al., 2005; Ullmann, 1985). Another possible explanation is the potential endogeneity problem between EP and FP. As indicated by the slack resource theory, successful firms tend to have more slack resources and can invest more in EM (McGuire et al., 1990; Waddock and Graves, 1997). Also, good management theory indicates a high correlation between good management practice and EP because of the good relationships with various stakeholders with improved morale, productivity and satisfaction (Waddock and Graves, 1997).

All these different opinions and conflicting findings indicate the complicated specifications and relationships between EP and FP (Hull and Rothenberg, 2008; Russo and Fouts, 1997). The lack of consensus conclusions fails to provide useful guidelines for business managers when deciding to invest in EM strategies and initiatives. Therefore, it is urgent and necessary for researchers to solve this problem by using more advanced and comprehensive analyses with better datasets and methodologies. This is where this study originates.

2. Purpose of study

This study conducts a systematic and comprehensive analysis by employing a large longitudinal dataset and multilevel methodology to test the relationship between EP and FP. The goal of this study is two–folded: the first is to reconcile the extant body of non–consensus results between EP and FP by identifying the heterogeneous effects across companies and industry sectors; the second is to figure out the causal relationship between EP and FP.

Specifically, different from most previous empirical studies that analyzed the static EP–FP relationship by using either cross–sectional data or aggregate analysis assuming homogeneous effect across companies and industry sectors, this study aims to decompose the "stable" and "aggregate" relationship between EP and FP to identify the variations in EP–FP relationship into different levels. We argue that the relationship may seldom be constant or stable for all individual companies under all conditions (Christmann, 2000; Russo and Fouts, 1997). Therefore, it will be fascinating to explore various scenarios under which a company's EP would have a positive, neutral or even negative impact on its FP. For example, as indicated by King and Lenox (2001), it is necessary to explore the roles of different firm characteristics in shaping the relationship between EP and FP. Moreover, industry characteristics and heterogeneities may also play a role (Elsayed and Paton, 2005), because companies within different industry sectors may face different requirements or standards in their operations. For example, the environmental compliance costs are usually larger for companies that operate within "dirty" industry sectors than those in "clean" sectors. Therefore, companies in "dirty" sectors may achieve greater cost savings and be capable of establishing environmental competitive advantage from proactive EM strategies of resource conservation or crisis prevention (Klassen and McLaughlin, 1996).

Meanwhile, studies that analyzed the causal direction from EP to FP are enormous, while research regarding the opposite causal direction is minimal. In most studies, the relationship between EP and FP was identified through cross–sectional analyses with only a stable snapshot of the association between EP and FP, which cannot exclude the possibility of reverse causality from FP to EP. Thereby it is necessary to address the potential issue of the temporal bi–directional causality between EP and FP using longitudinal data. Such data can provide a distinct advantage over cross–sectional data to identify causality (Hedström and Ylikoski, 2010). However, analyses in previous empirical studies of EM and sustainability are limited to aggregate regression models where a reciprocal relationship between EP and FP couldn't be estimated simultaneously. It is necessary to use advanced methodologies and good datasets to identify the full picture regarding the EP–FP relationship for a better understanding of the underlying rationales.

Therefore, data limitation, heterogeneities across firms and industries, endogeneity, as well as methodological limitation are possible reasons that lead to those mixed results regarding the EP–FP relationship in previous empirical studies. This study aims to solve those problems by building a multilevel framework to investigate the complicated EP–FP relationship. The objective of this study is to conduct a comprehensive analysis of the EP–FP relationship in order to reconcile previous non–consensus findings, identify different conditions under which it pays to be green or it doesn't pay to be green and investigate the existence of bi–directional causality between EP and FP.

3. Theoretical foundations and research framework

3.1. Dynamic capabilities

As the extension to the resource-based view (RBV) that recognizes the importance of unique and inimitable assets in creating competitive advantage (Teece, 1984; Wernerfelt, 1984), dynamic capabilities (DC) explains the mechanism that sustains the competitive advantage by arguing that the enduring source of a firm's competitive advantage comes from its capabilities of acquiring, integrating and deploying internal and external resources to match the market environment, rather than the possession of idiosyncratic resources (Teece et al., 1997; Lengnick-Hall & Wolff, 1999; Priem and Butler, 2001). An organization's dynamic capabilities enable its long-term competitive advantage and explain organizational adaptation to external dynamic environment (O'Reilly & Tushman 2008). DC emphasizes the dynamic process of resource acquisition, integration and deployment to fit with the external changing environment. DC has been widely used to explain intra-industry performance differences across firms (Zott, 2003). A firm's DC can effectively change its resource positions (Eisenhardt and Martin, 2000), capabilities (Kogut and Zander, 1992), and operational behaviors and efficiency (Nelson and Winter 1982). Therefore, to deal with the external pressure of sustainable development, firms vary in their resource configurations and capabilities when implementing environmental strategies for pollution/waste control and efficient use of resources. These variations will lead to different management capabilities, responsiveness to market, product design and innovation, operational costs, market position and therefore performance. Thus, the EP–FP relationship may vary across firms due to the variations of their DC.

3.2. Contingency theory

Contingency theory states that there is no universal design and the most effective structure designs are those that fit their related contingencies (Donaldson, 2001). Industry context has been recognized as an important contingency in determining performance (Barney and Ouchi, 1986; Hansen and Wernerfelt, 1989). Literature in strategic management and industrial organization has consistently emphasized the importance of industry underlying structural characteristics such as norms and practices in shaping a firm's performance. Different industries carry different environmental contingencies and structures. Datta et al. (2005) indicate that industry capital intensity, industry growth, and industry differentiation impact performance directly and indirectly. Organizations whose effective deals with and closely match their industrial context are more likely to improve performance than those who do not (Pfeffer, 1997). Thus, from the perspective of contingency theory, the EP–FP relationship may vary across industries because of the different industry configurations and structures.

3.3. Research framework

Based on the theoretical foundation discussed above, a multilevel framework is specified in this study to account for the across–firm, as well as across–industry variations. Moreover, bi–directional relationship with time sequence between EP and FP is also specified to address this potential problem of endogeneity. By using the multilevel framework, we aim to identify the following three hypotheses.

(1) Firm heterogeneity due to a firm's unique capabilities in resource acquisition, integration and deployment characteristics when implementing environmental strategies and pursuing financial performance. Based on DC, firms vary in their reaction to turbulent environment in responsiveness to markets, product design and innovation, managerial capabilities, etc.
(2) Industry heterogeneity due to variation of industrial configurations and structures. Base on contingency theory, higher performance can be achieved when an organization's configurations and behaviors closely match the associated industrial contexts and structures. For example, firms in green industries can probably achieve higher returns than those in dirty industries because of the lower costs in compliance with the requirements of environmental regulations (King and Lenox, 2001).
(3) Bi-directional relationship between EP and FP: A firm's EP may have a great impact on FP as shown in many previous studies. It will be equally likely to argue the opposite direction of causality with FP influencing EP (Wagner et al., 2001). EM initiatives are costly. The slack resource theory indicates that a firm with a high level of FP may have slack resources to invest in EM initiatives to improve its EP (McGuire et al., 1990; Waddock and Graves, 1997).

4. Methodology

4.1. Data

The data used in this study include the U.S. 500 largest publicly traded companies based on market capitalization headquartered in the U.S from 2005 to 2014, because large companies are more likely to adopt EM strategies in their operations (Ashby et al., 2012; Hart and Ahuja, 1996; Pagell and Wu, 2009). The data are consist of three parts from three different sources. The first part is regarding the measurement of EP, which is obtained from Bloomberg's database on Environmental, Social and Governance (ESG) for individual companies. The second part is about a company's financial performance and associated characteristics, like size, R&D, etc. obtained from Standard and Poor's COMPUSTAT database. Besides, the third part is industry characteristics including capital (structure and equipment) expenditures, collected from the U.S. Census Bureau. The dataset is consistent with our model with a three–level nested structure overall. The overtime repeated observations for an individual company (level 1) are nested in that specific company (level 2), and companies belonging to a specific industry sector are nested in that sector (level 3).

4.2. Measurement of EP

Measuring EP has been a difficult task due to the lack of consensus measurement instrument. Thus, subjective metrics and indicators were used in many previous studies, which is one of the primary reasons leading to inconsistent EP–FP relationships in literature. As environmental and social concerns are becoming increasingly important, many companies started to provide information and disclosures regarding their performance in environmental and other social aspects. More and more information regarding the measurement of a company's sustainable and ethical aspects is available from different sources, such as Bloomberg, Jantzi/Sustainalytics, MSCI/KLD Research and Analytics, Thomson–Reuters/Asset4, etc. Among these sources, Bloomberg, as a leading company in business and financial information, collects Environmental, Social and Governance (ESG) data from individual company's published reports or materials and evaluate each company's performance annually in these three aspects. The ESG dataset is integrated into Bloomberg's Equities, Intelligence, and Fixed Income platforms, and free for its user to use in the Bloomberg Equities and Industries service. For each company, its performance in environment, social and corporate governance is estimated between 0 and 100, based on the extent that the corresponding aspect is fulfilled in each associated company. Then a weighted-average ESG score is calculated based on a unique scheme with different weights in E, S, and G by considering the underlying risk factors for each industry and company.

The measurement of EP for a company used in this study is based on the environmental aspect of Bloomberg's ESG database. Environmental indicators used in Bloomberg's ESG database are related to energy, emissions, water, waste, and green operational policies, management, etc., such as include carbon emissions, energy efficiency, hazardous waste management, the use of recycled materials and natural resources, clean technology, etc. Although most information regarding ESG performance relies on a company's self-disclosure and third-party sources, “Bloomberg ESG rating will penalize companies for ‘missing data.’” (DavisPolk, 2017). Companies with good ESG performance normal choose to disclose their ESG activities extensively to generate favorable publicity for higher value, while firms with bad ESG performance would report minimally (Cahan et al., 2015; Fatemi et al., 2017). Therefore, the environmental score from Bloomberg's ESG database can represent a company's efforts in reducing irresponsible environmental activities and adopting proactive EM strategies for green operation and better environmental performance.

4.3. Measurement of FP and company characteristics

The data related to company's FP and characteristics are obtained from Standard and Poor's COMPUSTAT database that contains financial statement information on most publicly traded corporations in the U.S. This database provides a broad range of information for a company's financial status, including annual and quarterly income statements and related items in balance sheets (i.e. assets, liabilities, and equity), statement of cash flows, general company information and many other supplemental data items. The COMPUSTAT dataset was obtained from Wharton Research Data Services (WRDS).

Return on assets (ROA) will be used as a proxy for a company's FP. ROA is calculated as the ratio of net income to total assets, indicating a company's profitability based on its total assets. ROA demonstrates the operating efficiency in utilizing the company's total assets. ROA has been used by many studies in strategy, finance and accounting as well as corporate ethics and sustainability in the literature to measure a company's financial performance (Hart and Ahuja, 1996; King and Lenox, 2001; Nelling and Webb, 2009). Several other indicators from COMPUSTAT related to company characteristics will also be included in the dataset, such as firm size (i.e. the number of employees, assets), research and development (R&D) expense, capital expenditure, market shares, advertising expenses, etc. These indicators act as firm–level factors that may differentiate some companies from others in their impacts on the EP–FP relationship.

4.4. Measurement of industry characteristics

The industry–level data are obtained from the U.S. Census Bureau. The first one is capital expenditures for structures and equipment for each industry sector. Capital expenditure varies across industries. Large capital expenditures are required for industries (i.e., utilities, transportation, energy, etc.) that invest in expensive facilities, infrastructures and major manufacturing equipment. The second one is industry research and development expenditures. Spillover effect enables firms can take advantage of the accumulated pool of knowledge as well as innovation to obtain competitive advantages.[1] The third one is industry concentration measuring "the extent to which industry output is produced by a few firms and is a commonly used inverse proxy for industry competitiveness" (Melville et al., 2007). Firms in highly-concentrated industries tend to have bigger market capitalizations, which may affect their financial performance. Industry concentration is measured by using sales-based and asset-based Herfindahl indexes as indicated in Hou and Robinson (2006).

Overall, to satisfy the requirements of a multilevel longitudinal model, the data for each company need to have at least three–year observations continuously in order to estimate the cross–lagged multilevel models. Among those 500 largest U.S. companies from 2005 to 2014, 39 of them were removed from the dataset due to not satisfying this requirement. Thus, 461 companies were identified among this ten–year period with a minimum of 3 and maximum of 10 continuous observations. Companies are indicated with GVKEY.[2] The dataset includes 3668 company–year observations with an average of about 8 observations for each company. Overall, 38 industry sectors were identified (details provided in Appendix 1) based on the first three digits of these 461 companies' NAICS codes, with an average of about 12 companies in each sector, and a minimum of 5 and a maximum of 44 companies per sector.

5. Model setup

5.1. Null models for variance decomposition

As indicated by Raudenbush and Bryk (2002), a good starting point for any multilevel analysis is to use a null model for variance decomposition. To decompose the variances of EP and FP, the longitudinal dataset is fitted by using an unconditional means multilevel model which is a model without any predictor. It simply helps to partition the total variance of a dependent variable into different levels (Fitzmaurice et al., 2012; Hoffman, 2015; Singer and Willett, 2003). Models ***1a*** and ***1b*** are constructed to decompose the total variances of FP and EP correspondingly by using the nested–structure dataset described earlier.

Model ***1a***: Level 1: $$FP_{ijt} = \alpha_{0ij} + c_{ijt}, \tag{1}$$

Level 2: $$\alpha_{0ij} = \beta_{0j} + d_{ij}, \tag{2}$$

Level 3: $$\beta_{0j} = \theta_0 + e_j. \tag{3}$$

Here, level 1 captures the within–company variation over time. FP_{ijt} is the FP of company i in industry sector j at time t. α_{0ij} is the over–time mean of FP for company i in industry j. c_{ijt} is the residual for company i at time t, indicating company i's annual FP deviation from its over–time mean. Level 2 captures the between–company variation with β_{0j} as the mean FP for all companies in industry j. d_{ij} is residual, indicating company i's FP deviation from the industry–level mean, β_{0j}, in industry j. Level 3 captures the between –industry variation with θ_0 as the grand mean of FP for all companies and e_j as the corresponding residual for the FP deviation of industry j. All the residuals (i.e., c_{ijt}, d_{ij}, e_j) at different levels are assumed to follow normal distributions.

Model ***1a*** can be rewritten as follows by combining the above three–level equations into one equation,

$$FP_{ijt} = \theta_0 + e_j + d_{ij} + c_{ijt}. \tag{4}$$

Here, e_j captures the random effect of industry j, d_{ij} can be viewed as the random effect of company i within industry j, and c_{ijt} is the overtime variation. This model is also regarded as *null random intercept model* because of their intercepts, $\theta_0 + e_j + d_{ij}$, are not constants but variables following normal distributions. With this decomposition, the total variance of FP, i.e. $Var(FP_{ijt})$, can be decomposed into three components, between–industry variance, $Var(e_j)$, between–company variance, $Var(d_{ij})$, and within–company variance over time, $Var(c_{ijt})$ as follows,

$$Var(FP_{ijt}) = Var(e_j) + Var(d_{ij}) + Var(c_{ijt}). \tag{5}$$

Similarly, model ***1b*** is used to decompose the variance of EP as follows,

Model ***1b***: Level 1: $$EP_{ijt} = \gamma_{0ij} + l_{ijt}, \tag{6}$$

Level 2: $$\gamma_{0ij} = \eta_{0j} + m_{ij}, \tag{7}$$

Level 3: $$\eta_{0j} = \kappa_0 + n_j. \tag{8}$$

Here, EP_{ijt} is the observed EP for company i in industry sector j at time t. Level 1 captures the within–company changes over time with γ_{0ij} as over–time mean EP of company i in industry j. l_{ijt} is residual at time t. Level 2 captures the between–company variation with η_{0j} as the mean EP of companies in industry j. m_{ij} is the residual for company i's EP deviation from the industry mean, η_{0j}. Level 3 captures the between–industry variation with κ_0 as the grand mean of EP and n_j as the residual for industry j's EP deviation from grand mean κ_0.

Model ***1b*** can also be rewritten as follows by combining the above three–level equations into one equation for a *null random intercept model.*

$$EP_{ijt} = \kappa_0 + n_j + m_{ij} + l_{ijt}. \tag{9}$$

n_j represents the random effect of industry j, m_{ij} represents the random effect of company i within industry j, and l_{ijt} captures the overtime variation. All the residuals (i.e., l_{ijt}, m_{ij}, n_j) are assumed following normal distributions. The total variance of EP can also be decomposed into three parts, industry-level variance $Var(n_j)$, company-level variance $Var(m_{ij})$ and over-time variance $Var(l_{ijt})$ as follows,

$$Var(EP_{ijt}) = Var(n_j) + Var(m_{ij}) + Var(l_{ijt}). \tag{10}$$

5.1.1. Relationship between FP and EP

Most studies about the EP–FP relation have mainly focused on the impact of EP on FP. For a full consideration of the EP–FP relationship, it is important and necessary to study the potential influence of FP on EP as well. This is a question of endogeneity that whether successful firms have more resources to spend on EM initiatives or whether better EP leads to better financial outcomes afterward. If an endogeneity problem exists, many previous studies focusing on the EP→FP causality may suffer from it with biased outcomes in their analyses. Based on the virtuous circle perspective, the EP–FP relationship may be bi–directional and involve iterative impact processes in a time–ordered recursive sequence between EP and FP (Orlitzky et al., 2003). That is to say, a firm's prior FP has an impact on its subsequent EP which will, in turn, influence its FP in the following period (Orlitzky et al., 2003).

To address this potential endogeneity problem regarding the two–way interaction between EP and FP, simultaneous equation models are used to estimate the bi–directional relationship with time sequence specified between EP and FP. An autoregressive cross–lagged panel models (ACLPM) (Cole and Maxwell, 2003; Curran, 2000) is constructed to simultaneously address the reciprocal effects between EP

[1] The author truly appreciates one anonymous reviewer pointing this out.

[2] GVKEY is a unique 6-digit number assigned to each company in the COMPUSTAT database. It is a company identifier and represents the primary key for a company as an index constituent.

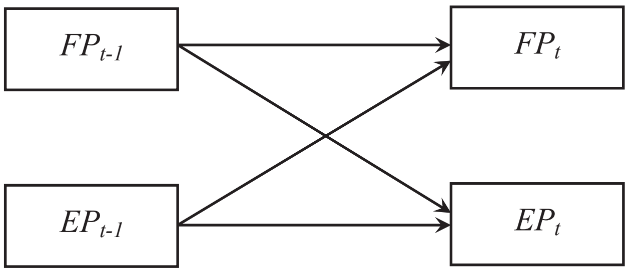

Fig. 1. Temporal cross–lagged model.

and FP. The EP–FP relationship involves temporal iterative processes of the impact that link the two variables together in a *time–ordered recursive sequence*, as shown in Fig. 1 below.

In an ACLPM, variables at time (t) account for deviations at a previous time (t-1). Autoregressive effect is estimated by regressing a variable from a previous time point to the next time point, while the cross-lagged effect is calculated by regressing the one variable at time (t) on the other variables' previous time (t-1). The temporal order of one variable before another can help to identify the causal relationship. In Fig. 1 above, the single–headed arrows from FP_{t-1} to FP_t, and EP_{t-1} to EP_t represent the autoregressive effects between the same variable at different occasions. The diagonal single–headed arrows from FP_{t-1} to EP_t, and EP_{t-1} to FP_t represent the reciprocal EP–FP relationships of impact over time.

Specifically, fixed–slope and random–slope ACLPM will be used to estimate, (1) the aggregate overall EP–FP relationship, and (2) the disaggregate EP–FP relationship at the company level and industry level.

Model 2***a***: random–intercept and fixed slope model,

$$FP_{ijt} = \theta_0 + e_j + d_{ij} + \theta_1 FP_{ij(t-1)} + \theta_2 EP_{ij(t-1)} + \boldsymbol{\beta} \boldsymbol{X}_{ij(t-1)} + \varepsilon_{ijt}, \tag{11}$$

$$EP_{ijt} = \kappa_0 + n_j + m_{ij} + \kappa_1 EP_{ij(t-1)} + \kappa_2 FP_{ij(t-1)} + \gamma \boldsymbol{X}_{ij(t-1)} + £_{ijt}. \tag{12}$$

Here, the fixed-slope model identifies the overall relationship between EP and FP by specifying the impact of EP on FP in equation (11) and FP on EP in equation (12), with some control variables.

In equation (11), e_j can be viewed as the random effect of industry *j* on FP, and d_{ij} can be viewed as the random effect of company *i* within industry *j* on FP. θ_1 is the autoregressive coefficient indicating the effect of FP on itself measured at a later time point. A small or zero autoregressive coefficient means a substantial change of FP over time, while a significantly positive autoregressive coefficient means that a company's FP is relatively stable over time. θ_2 is the coefficient of EP on FP in the next period, representing the overall mean impact of EP on FP for all companies in the dataset. $\boldsymbol{\beta}$ is a row vector of coefficients of the column vector of control variables, $\boldsymbol{X}_{ij(t-1)}$, for company *i* in sector *j*, indicating the average effects of control variables on FP for all companies.

Similarly, in equation (12), n_j represents the random effect of industry *j* on EP, and m_{ij} represents the random effect of company *i* within industry *j* on EP. κ_1 is also an autoregressive coefficient indicating the effect of EP on itself measured at a later time point. κ_2 is the coefficient of FP on EP in the next period, representing the overall mean impact of FP on EP for all companies included in the dataset. γ is a row vector of coefficients for the column vector of control variables, $\boldsymbol{X}_{ij(t-1)}$ of all companies, measuring the average effect of those control variables on EP.

Both θ_2 and κ_2 describe the cross–lagged mean effects between FP and EP, which are the average impacts of all companies' FP on their subsequent EP and the effect of their EP on subsequent FP. A significantly positive value of θ_2 indicates the mean impact of EP on following–period FP is positive, and a significantly negative value of θ_2 indicates a negative impact. Similarly, a significantly positive value of κ_2 means that an increase in FP can lead to enhanced EP in the following period, and vice versa. Similar to many previous studies, model 2***a*** is an aggregate analysis by studying the average relationship between EP and FP.

Model 2***b***: random–intercept and random–slope model at both company and industry levels,

$$FP_{ijt} = \theta_0 + e_j + d_{ij} + \theta_1 FP_{ij(t-1)} + (\theta_2 + p_j + s_{ij}) EP_{ij(t-1)} + \boldsymbol{\beta} \boldsymbol{X}_{ij(t-1)} + \varepsilon_{ijt}, \tag{13}$$

$$EP_{ijt} = \kappa_0 + n_j + m_{ij} + \kappa_1 EP_{ij(t-1)} + (\kappa_2 + q_j + w_{ij}) FP_{ij(t-1)} + \gamma \boldsymbol{X}_{ij(t-1)} + £_{ijt}. \tag{14}$$

Compared with model ***2a***, the random-slope model ***2b*** estimates the specific relationship between EP and FP for each specific company and industry included in this dataset.

Specifically, p_j in equation (13) and q_j in equation (14) represent the random part (i.e., deviation from the grand mean effects θ_2 and κ_2) of the cross–lagged effect in industry *j* between FP and EP by allowing this EP–FP relationship to vary across industry sectors. That is, the cross–-lagged impacts between FP and EP may be different across industry sectors if there is any significant random industry effect. s_{ij} equation (13) and w_{ij} in equation (14) are also incorporated in model ***b*** to capture the random company effect on the cross–lagged relationship between FP and EP for company *i* in industry *j*, with the goal of finding the unique EP–FP relationship for company *i* in industry sector *j*. Significant deviations of s_{ij} and w_{ij} from grand mean values θ_2 and κ_2 also indicate the existence of significant variations in the FP–EP relationship across companies in industries *j*. p_j, q_j, s_{ij} and w_{ij} are all assumed following normal distributions. θ_2 and κ_2 in model ***b*** are means of cross–lagged effects between FP and EP for all companies in all industry sectors.

For example, in equation (13), the impact of EP on FP for a specific company *i* in industry *j* is determined by the mean effect θ_2 for all companies in all sectors as well as the random industry effect (p_j) and the random company effect (s_{ij}). That is, the total impact of EP on FP for company *i* in industry *j* is determined by $\theta_2 + p_j + s_{ij}$. Under some circumstances, even though the overall mean effect (i.e., θ_2) is positive as indicated in many previous studies, the total impact of EP on FP is not necessarily positive for a specific company in certain industry sectors if an negative random effect at industry or company level exists. Thus zero or even negative association could be possible for some companies in certain industries. These arguments also apply to equation (14) when studying the impact of FP on subsequent EP. The impact of a company's FP on its subsequent EP includes the mean effect of FP on EP (κ_2), the industry effect (q_j) and company effect (w_{ij}). The total impact of FP on subsequent EP for a specific company *i* in industry *j* is given as $\kappa_2 + q_j + w_{ij}$, which is determined by the average effect κ_2, as well as the variations of company and industry effects, q_j and w_{ij}.

6. Results and discussion

6.1. Results of variance decomposition

As discussed in section 3.2.1, we use an unconditional means multilevel model is built to decompose the total variance of FP and FP into industry-level, company-level and over–time variances as displayed in Tables 1 and 2 below.

Based on Table 1, the total variance of the FP for all companies in the dataset is significantly decomposed into three levels, within company (5.20e–06), across companies (8.65e–05) and across sectors (1.65e–05), with corresponding standard errors in parentheses. This is the estimation results for model ***1a***. The variance partition coefficients (VPCs)[3] for FP at each level are 4.81%, 79.94% and 15.25% respectively. The within–company variance only accounts for 4.81% of the total variance, indicating that the FP of all companies in the dataset is relatively stable over time. The across–company variance accounts for 79.94% of total FP variance, which means that FP varies a lot across companies. Industry –level variance account for 15.25% of the total FP

[3] VPCs are used to calculate the relative magnitude of the variance components at each level to the total variance. VPCs report the proportion of response variance at each level.

International Journal of Production Economics 219 (2020) 328–340

Table 1
Decomposition of FP variance.

DV: FP	Variances *equation (4)*	Variance Partition Coefficients (VPCs)	Intraclass Correlation Coefficients (ICCs)
Industry Level, $Var(e_i)$	1.65e–05 (5.84e–06)[a]	0.152	0.152
Company level, $Var(d_{ij})$	8.65e–05 (6.02e–06)[a]	0.799	0.952
within company, $Var(c_{ijt})$	5.20e–06 (2.00e–07)[a]	0.048	0

[a] Indicates significance with p–value ≤ 5%.

Table 2
Decomposition of EP variance.

DV: EP	Three–Level Model *equation (9)*	Variance Partition Coefficients (VPCs)	Intraclass Correlation Coefficients (ICCs)
Industry Level, $Var(l_i)$	39.600 (13.111)[a]	0.177	0.177
Company level, $Var(m_{ij})$	156.376 (11.227)[a]	0.699	0.876
Within Company, $Var(n_{ijt})$	27.868 (1.077)[a]	0.124	0

[a] Indicates significance with p–value ≤ 5%.

variance, indicating variations across industries. Moreover, the intraclass correlation coefficients (ICCs)[4] of FP at each level are also calculated. The industry ICC is 0.152, indicating that the correlation between two FP observations of two different companies in the same industry is 0.152. The company ICC is 0.952, indicating the correlation of two FP observations overtime within a single company is 0.952. Therefore, a company's FP overtime is very stable, which is consistent with the results of VPCs.

Similarly, the results of EP shown in Table 2 follow a very similar pattern as FP. The total variance of company's EP is significantly decomposed into three levels, within company (27.868), across company (156.376) and across sectors (39.600), with corresponding standard errors in parentheses. The VPCs for EP at each level are calculated as 12.45%, 69.86% and 17.77% respectively. The industry ICC is 0.152 indicating the correlation between two EP observations of two different companies in the same industry is 0.152. The company ICC is 0.876, indicating a high correlation between two EP observations overtime within a single company.

Based on the results of VPC and ICC, for both FP and EP, there are significant variances across industries and companies, and there are great similarities for companies within the same industry and for overtime observations within the same company. Therefore, a multilevel model is appropriate to analyze the dataset used in this study.

6.2. Results for FP

Based on the results in Table 3 (DV: FP) above, the autoregressive coefficient of FP, θ_1, is significantly large, 0.708 (0.0144, $p < 0.01$) for fixed–slope model and 0.707 (0.0139, $p < 0.01$) for random–slope model, which indicates high stability for company's FP over time. For the overall relationship between EP and FP, on average, a company's EP does have a positive impact on its subsequent FP, indicated by the significantly positive coefficient of EP on FP in fixed slope model (*equation (1)*) with $\theta_2 = 0.0000765$ ($p < 0.01$). That means, by considering all the companies in the dataset together, an increase in a company's EP can improve its FP in the following period. EP and FP are positively related, which is consistent with many previous studies that concluded the positive EP–FP relationship.

The random–slope model specified by *equation (3)* decomposes the EP–FP relationship into company level and industry level. The significant variance of p_j, $Var(p_j)$, indicates significant variations in the impact of EP on FP across industries. And the significant variance of s_{ij}, $Var(s_{ij})$, indicates significant variations in the impact of EP on FP across companies. Thus, even though the overall mean impact of EP on FP is significantly positive with $\theta_2 = 0.0000685$ ($p < 0.01$) as indicated in the random–-slope model, *equation (3)*, EP may influence FP differently from this average impact of θ_2 for some companies in certain industry sectors.

6.2.1. Random industry effect

The variations of the random effects for the impact of EP on FP at industry level are displayed in the "caterpillar plot" below (Fig. 2) with the estimated values for the magnitudes of random effects for each industry sector and associated errors in Appendi × 1.

Fig. 2 shows the estimated magnitudes of variations regarding the impact of EP on FP across 38 industry sectors included in the dataset. Based on the results in Appendix 2 and the plot in Fig. 2 below, it can be found that although the random industry effect (i.e. p_j) for most industry sectors is not significantly different from 0 with 95% confidence intervals covering the value of 0, there are several industry sectors with significant random impacts of EP on FP (i.e. p_j) above as well as below 0. These variations indicate that for some industry sectors, the financial return from investment in EM strategies is below the average with negative random industry effect ($p_j < 0$), while some other industry sectors have above-average return from EM strategies with positive random industry effect ($p_j > 0$). From the results, we can find that there are three industries that have negative random industry effects with one in utilities and two in manufacturing (food and fabricated metal product). This below-average return from EM activities compared with other industries possibly due to the lack of public acknowledgment or poor industry structure. The two industries with above-average random effect are "Petroleum and Coal Products Manufacturing" and "Credit Intermediation and Related Activities". It is interesting that the public will reward the EM efforts in the industry of Petroleum and Coal Products Manufacturing, indicating that society is paying more and more attention on sustainability.

Specifically, for those companies belonging to industry j, the impact of their EP on FP is defined as $\theta_2 + p_j$ if only random industry effect (p_j) is considered. Even though the average effect θ_2 is positive, $\theta_2 + p_j$ might be close to zero or even negative if p_j is negative. For some industry sectors with positive values of p_j, the impact of EP on FP ($\theta_2 + p_j$) would be higher than average.

For example, for the industry effect of sector 221 in appendix 2, sector of utilities, p_j is estimated as -0.000125 (0.000045, $p < 0.01$). This significantly negative random industry effect will decrease the overall impact of EP on FP for companies in the utility sector. Because the mean impact of EP on FP in *equation (3)* θ_2 is estimated as 0.0000685, the result of $\theta_2 + p_j$ will be given as -0.00056, indicating a mean negative impact of EP on FP for companies in industry sector utilities. This negative impact indicates that the financial return from EM activities for companies in the utility sector might not cover the

[4] ICCs measure the homogeneity of the observed values within a given group. High ICC values indicate high similarity among observed units within a cluster.

Table 3
Impact on FP in ACLPM for equations (1) and (3).

DV: FP_t		Fixed slope *equation (1)*	Random slope *equation (3)*	Random slope *for X*
Coefficient of FP_{t-1} (θ_1)		0.708 (0.0144)[a]	0.707 (0.0139)[a]	0.705 (0.0163)[a]
EP_{t-1}	Coefficient (θ_2)	7.65e–05 (1.02e–05)[a]	6.85e–05 (9.40e–06)[a]	6.61e–05 (9.65e–06)[a]
	$Var(p_j)$	n/a	8.02e–09 (3.25e–09)[a]	7.87e–09 (3.66e–09)[a]
	$Var(s_{ij})$	n/a	1.12e–08 (3.25e–09)[a]	1.22e–08 (3.41e–09)[a]
RD_{t-1}	Coefficient (β_1)	3.03e–07 (8.76e–08)[a]	3.58e–07 (8.72e–08)[a]	4.79e–07 (9.82e–08)[a]
	$Var(w_{1ij})$	n/a	n/a	6.67e–13 (4.75e–13)
EMP_{t-1}	Coefficient (β_2)	3.29e–06 (7.68e–07)[a]	2.97e–06 (7.32e–07)[a]	1.43e–05 (2.02e–06)[a]
	$Var(w_{2ij})$	n/a	n/a	1.67e–09 (5.78e–10)[a]
CE_{t-1}	Coefficient (β_3)	2.72e–07 (5.38e–07)	2.68e–07 (6.05e–07)	2.61e–07 (6.15e–07)
	$Var(w_{3ij})$	n/a	n/a	3.72e–13 (1.55e–13)
MS_{t-1}	Coefficient (β_4)	0.082 (0.027)[a]	0.081 (0.025)[a]	0.084 (0.028)[a]
	$Var(w_{4ij})$	n/a	n/a	0.019 (0.0089)[a]
AD_{t-1}	Coefficient (β_5)	4.78e–07 (5.84e–07)	4.82e–07 (6.01e–07)	4.76e–07 (6.17e–07)
	$Var(w_{5ij})$	n/a	n/a	1.70e–09 (6.04e–10)[a]
$ICON_{t-1}$	Coefficient (β_6)	0.0073 (0.0026)[a]	0.0067 (0.0031)[a]	0.0069 (0.0028)[a]
	$Var(w_{6j})$	n/a	n/a	0.00129 (0.00037)[a]
IRD_{t-1}	Coefficient (β_7)	5.52e–07 (7.82e–08)[a]	5.65e–07 (7.61e–08)[a]	5.32e–05 (2.17e–06)[a]
	$Var(w_{7j})$	n/a	n/a	7.17e–06 (2.18e–06)[a]
ICA_{t-1}	Coefficient (β_8)	1.72e–07 (3.05e–08)[a]	1.68e–07 (3.32e–08)[a]	1.64e–07 (3.05e–08)[a]
	$Var(w_{8j})$	n/a	n/a	4.78e–14 (2.57e–14)[a]

[a] Indicates significance with p–value ≤ 5%.

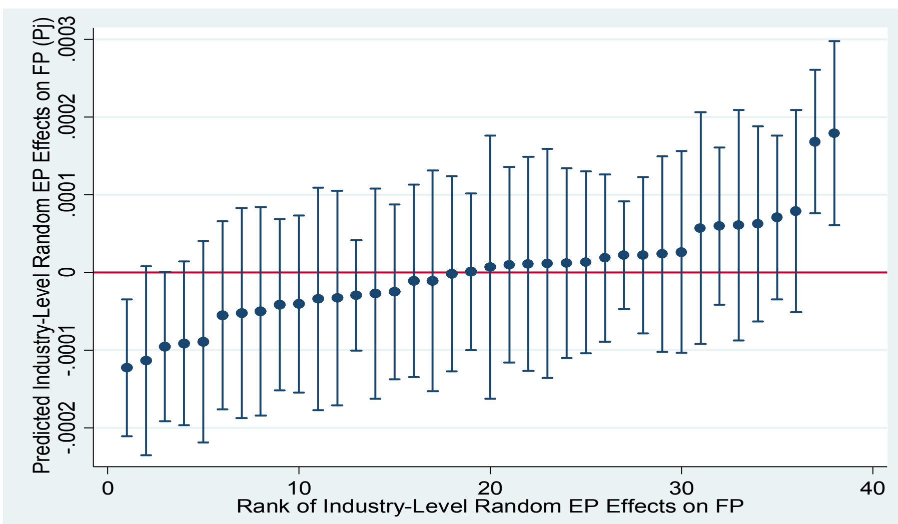

Fig. 2. Predicted magnitudes of industry effects on FP.

associated costs of EM activities. One possible reason of this negative EP–FP relationship could be that customers on the market do not fully trust and value the declaration or efforts of EM initiatives from companies operating in "dirty" industry sector like utilities. Another possible reason could be that the facilities and equipment required by the EM initiatives are so expensive, which cannot be fully covered by the financial return from EM in short-term (King and Lenox, 2001). Therefore, to merely improve FP in short or intermediate term, it may be worthwhile for them to reduce their expenditures in EM activities based on the results in this study. Therefore, the economic choice for those companies operating in industries with a negative EP's impact on FP could be simply following the minimum of environmental regulatory requirements without doing any extra EM initiative beyond.

6.2.2. Random company effect

Similar to the random effect of EP on FP at the industry level, the impact of EP on FP also varies across companies indicated by the significant variance of the random company effect, $Var(s_{ij})$, in *equation (3)*. The "caterpillar plot" of Fig. 3 below shows the estimated magnitudes of variations regarding the impact of EP on FP across all 461 companies.

Based on the plot in Fig. 3 above and the results in appendix 3, most of the estimated random company effects (s_{ij}) for the 461 are not significantly different from zero with 95% confidence intervals. Only five of them are significantly different from zero with three negative and two positive. One possible reason is that we only include large companies in our dataset. Large companies tend to have better management capability and more easily to achieve economy of scale regarding their EM strategies than smaller companies. Larger variation in company's random effects could be possible if small companies were included. Significantly negative company effect, s_{ij}, indicates that the associated company has below-average financial return from its EM investments, probably due to their inefficient management capabilities or other constraints inside the company. On the other hand, significantly positive company effect of s_{ij}, indicates that the associated company has above-average financial return from its EM investments, possibly because of efficient management or high green reputation.

If only random company effect is considered, the impact of a company's EP on FP is specified as $\theta_2 + s_{ij}$ that could be close to zero or even negative if s_{ij} is significantly negative even though θ_2 is positive. For example, from the results in appendix 3, the estimated random company effect for company 027786 (GVKEY) is given as $s_{ij} = -0.0002072$ (0.0000863, $p < 0.01$), This significantly negative

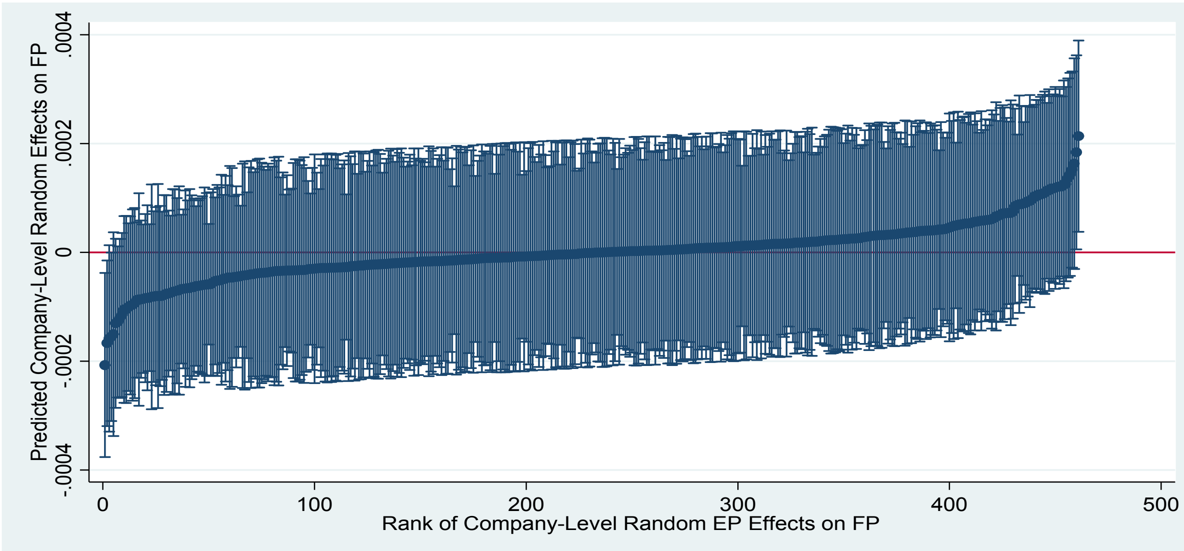

Fig. 3. Predicted magnitudes of company effects on FP.

company random effect will lead to a negative effect EP on FP, as $\theta_2 + s_{ij} = -0.0001387$, which means that the FP of company 027786 will decrease as its EM expenditures increases. Therefore, in short run, it might be a good economic choice for this company just following the minimum of regulatory requirements without implementing extra EM activities beyond the requirements.

To summarize the impact of EP on FP, a positive overall relationship has been found in this study, which is consistent with many of previous studies that proposed a positive association between EP and FP. However, the decomposition of EP–FP relationship leads to negative results for some companies in certain industry sectors probably due to their unique characteristics, which worth further investigation in future studies.

6.2.3. *Effects of control variables*

Several company-level and one industry-level control variables are also included in the ACLPM model. The first one is a company's expenditures in research and development (R&D), representing the company's innovation capacity, indicated by *RD* in Tables 3 and 4. R&D has positive impact on FP with coefficient β_1 = 3.03e–07 (8.76e–08, $p < 0.01$) specified by equation (1), and β_1 = 3.58e–07 (8.72e–08, $p < 0.01$) in equation (3). The second control variable is company size by using the number of employees, *EMP*. The significantly positive coefficients of EMP in *equation (1)* and *(3)*, β_2, indicate that larger companies tend to have higher FP. The third control variable is company's capital expenditure (*CE*) has insignificant effect on FP. The fourth one is market share (*MS*) that has significant effect on FP. The fifth one is company's advertising expenditure (*AD*) with no significant effect on FP. Three industry–level factors are also included in this study, the first one is industry concentration (*ICON*) industry–level total capital expenditures, denoted by IC. The significantly positive coefficients of *ICON*, β_6, in equations (1) and (3) suggest high–level FP for companies that operate in highly-concentrated industry sectors. The second and third industry-level variables are industry R&D and capital expenditure. Both have positive impact on company's FP. The possible reason for this could be the spillover effect resulted from the industry–level accumulated pool of capital and knowledge.

Lastly, for comparison purpose, the random effects for control variables are also estimated with results shown in the last column of Table 3. Although there are some variations for the impact of some of these control variables on FP across companies and industries, the results for the impact of EP on FP are pretty consistent with the previous fixed and random slope models. The impact of R&D on FP is homogeneous across companies with the variance of its random coefficient, $Var(w_{1ij})$, not significant. But the impact of company size (*EMP*) and market share (*MS*) vary across companies with significant variances, $Var(w_{2ij})$ and $Var(w_{4ij})$. At the industry level, the impact of industry concentration (*ICON*), R&D (*RD*) and capital expenditure (*ICA*) vary across industries with significant variances, $Var(w_{6j})$, $Var(w_{7j})$ and $Var(w_{8j})$. Since the impacts of the control variables are not the focus of this study, details regarding the random effects for those impacts are not discussed here.

6.3. *Results for EP*

Based on Table 4 (DV: EP) above, the autoregressive coefficient of EP, κ_1, is significantly large with a positive value of 0.786 (p $<$ 0.01) for both fixed–slope and random–slope models, indicating high stability of company's EP over time. Overall, a company's FP does have a positive impact on its EP, indicated by the significantly positive coefficient of FP on EP with κ_2 = 70.89 (28.56, $p < 0.01$). This result indicates that an increase in a company's FP can improve its EP in the following period. A company with good FP has extra funding to support its EM initiatives and efforts, which confirms some arguments in previous literature regarding the existence of reverse or bi–directional causality between EP and FP (Goss and Roberts, 2011; Wagner et al., 2001). However, there is not much variation regarding the impact of FP on EP across companies and industries, indicated by the insignificant variances of random effects, i.e. $Var(w_{ij})$ and $Var(q_j)$, in equations (2) and (4). One possible reason is that only large companies are included in our dataset. Large companies tend to be financially successful with more resources and high-level management capabilities. Large firms are more likely to invest more in EM initiatives to attract more public attention than small firms (Ashby et al., 2012). For the company- and industry-level control variables, only industry-level R&D has a significant impact on EP. This impact varies across industries possibly due to the spillover effect resulted from the industry–level accumulated pool of capital and knowledge.

Therefore, based on the results of ACLPM, bi–directional causality between EP and FP in a temporal sequence is confirmed as proposed in the previous literature (Goss and Roberts, 2011). For many previous studies that only investigated the one–direction impact of EP on FP, endogeneity problem might exist due to this reciprocally temporal relationship between EP and FP.

7. Conclusions, limitations, and future research

7.1. *Conclusions*

The results found in this study demonstrate variations regarding the

Table 4
Impact on EP in ACLPM for equations (2) and (4).

DV: EP_t		Fixed slope equation (2)	Random slope equation (4)	Random slope X
Coefficient of EP_{t-1} (κ_1)		0.786 (0.0147)[a]	0.782 (0.0149)[a]	0.779 (0.0142)[a]
FP_{t-1}	Coefficient (κ_2)	70.89 (28.56)[a]	70.77 (27.69)[a]	70.87 (27.82)[a]
	$Var(q_j)$	n/a	4.19e–12 (5.67e–09)	4.39e–12 (6.47e–09)
	$Var(w_{ij})$	n/a	1.13e–18 (7.27e–12)	1.12e–18 (7.11e–12)
RD_{t-1}	Coefficient (γ_1)	0.00015 (0.00016)	0.00015 (0.00016)	0.00016 (0.00016)
	$Var(f_{1ij})$	n/a	n/a	4.12e–23 (1.69e–22)
EMP_{t-1}	Coefficient (γ_2)	−0.00033 (0.0016)	−0.00034 (0.0015)	−0.00035 (0.0016)
	$Var(f_{2ij})$	n/a	n/a	2.66e–19 (3.82e–16)
CE_{t-1}	Coefficient (γ_3)	0.00078 (0.0016)	0.00077 (0.0018)	0.00078 (0.0016)
	$Var(f_{3j})$	n/a	n/a	1.07e–7 (1.69e–7)
MS_{t-1}	Coefficient (γ_4)	68.32 (75.65)	68.34 (75.64)	68.29 (75.75)
	$Var(f_{4ij})$	n/a	n/a	4.14e–23 (1.67e–22)
AD_{t-1}	Coefficient (γ_5)	−0.0048 (0.017)	−0.0048 (0.017)	−0.0048 (0.017)
	$Var(f_{5ij})$	n/a	n/a	3.45e–6 (3.78e–6)
$ICON_{t-1}$	Coefficient (γ_6)	18.31 (27.58)	18.29 (27.46)	18.30 (27.55)
	$Var(f_{6j})$	n/a	n/a	1.32 (1.58)
IRD_{t-1}	Coefficient (γ_7)	0.043 (0.019)[a]	0.044 (0.019)[a]	0.043 (0.019)[a]
	$Var(f_{7ij})$	n/a	n/a	0.024 (0.0058)[a]
ICA_{t-1}	Coefficient (γ_8)	8.12e–07 (1.9e–05)	8.12e–07 (1.9e–05)	8.13e–07 (1.9e–05)
	$Var(f_{8j})$	n/a	n/a	1.09e–21 (1.77e–20)

[a] Indicates significance with p–value ≤ 5%.

relationship between EP and FP across industry sectors and companies. Consistent with many previous studies, a mean positive impact is found between EP and FP for all companies included in our dataset. Moreover, heterogeneities regarding the impact of EP on FP is also found across companies and industry sectors by using multilevel analysis. Under certain scenarios for some companies and industry sectors, positive or negative impacts between EP and FP can be identified based on the data in this study, which reconciles those mixed findings in many previous empirical studies. Lastly, bi–directional causal relationship between EP and FP is found. The simultaneous estimation of the reciprocally temporal relationship between EP and FP solves the endogeneity problem concerned by some researchers (Garcia–Castro et al., 2010; Ghoul et al., 2011; Goss and Roberts, 2011).

The findings in this study indicate that the company's EP and FP are positively related overall, suggesting that in general a proactive EM strategy is helpful in improving future FP (i.e., it does pay to be green). However, not all companies in all industry sectors can just mimic that strategy and benefit from it. The heterogeneities of EP–FP associations under various company–level and industry–level contexts indicate that the actual impact of a company's EP on FP can be different from the aggregate effect and largely depends on its own characteristics and the industry sector it belongs to. The bi–directional causal relationship between EP and FP also implies that it is necessary for companies to have sufficient financial resources in order to implement proactive environmental strategies and initiatives.

This study contributes to the literature and in–world practitioners in the following five aspects. First, this study employs a multilevel framework to decompose the EP–FP relationship into different levels and finds that the relationship is not homogeneous for all companies and industries but varies under different contexts with different characteristics at the company level and industry level. A single explicit conclusion cannot be drawn simply regarding the relationship between EP and FP without providing certain conditions or contexts (Christmann, 2000; King and Lenox, 2001; Russo and Fouts, 1997). The findings in this study reconcile those non–consensus EP-FP relationships in previous studies by specifying positive as well as negative relationships between EP and FP under different contexts. Second, this study addresses the issue of endogeneity by constructing an autoregressive cross–lagged model and using longitudinal data. A bi–directional causality is found confirming the existence of iterative processes of impact between EP and FP in a time–ordered recursive sequence. EP and FP impact each other over time. A company's FP has an impact on its subsequent EP which in turn will impact on its FP in the following period (Orlitzky et al., 2003). Because EM initiatives are costly, a positive impact of company's FP on subsequent EP indicates that high–-level FP will provide necessary funding and resources for companies in helping their implementation of EM strategies and initiatives (McGuire et al., 1990; Waddock and Graves, 1997).

Thirdly, this study proposes a multilevel framework to analyze the EP–FP relationship. This multilevel aspect provides a promising direction for empirical research when employing data with various contexts or scenarios. Multilevel analysis can decompose the variance of observations into different levels across subjects and their upper levels. Independence assumption required by OLS can be relaxed in multilevel models by allowing lower–level observations nested in single higher–level subjects (Hoffman, 2015). Moreover, multilevel modeling can handle and estimate datasets with unbalanced structure and address the problem of potential collinearity (Browne et al., 2001; Patterson, 2013; Raudenbush and Bryk, 2002). Compared with pooled OLS and many other static analyses, multilevel analysis can identify the within– and between–subject heterogeneities regarding the relationship among variables interested. Therefore, multilevel modeling can aid the researchers to obtain a full picture of those complicated relationships in the real world.

Fourth, this study has practical implications for managers regarding their decisions in adopting EM strategies and investors with concerns of sustainability. Because EM initiatives are expensive with high uncertainty in improving FP (Figge et al., 2002; Hansen et al., 2007), the findings from this study are helpful for managers when directing resources and investments to improve EP. For business managers, their goal is to integrate environmental strategies into their decision makings and ensure sustainable financial success. By recognizing the heterogeneous EP–FP relationships across companies and industries, the managers can properly manage their EM activities based on their specific situations and contexts, rather than merely mimic others and increase investment in EM initiatives. For example, for companies with above–average impact of EP on FP at company as well as industry level, they can adopt proactive EM strategies and increase EP to not only meet the obligations of environmental regulations but also move beyond the regulations to obtain first–mover advantages and achieve the above–-average financial return. On the other hand, for managers whose companies have below–average effect of EP on FP either at company or industry level, they cannot mimic proactive EM strategies adopted by those with above-average impact. Instead, they need to manage their EM activities properly to just meet the requirements of environmental regulations and increase their management capabilities to enhance the efficiency of EM activities and innovations. For investors who concern environmental sustainability and plan to invest in certain companies or industries with the goal to maximize the return of their investment, it would be a better choice for them to direct their investment into those companies with above–average impact of EP on FP at both company and industry levels to gain above–average returns and achieve the win–win situation under which both their financial return and environment will benefit.

Lastly, this study also has some important implications for policymakers. Realizing the non–unanimous EP–FP relationships across companies and industries, policymakers need to design more flexible and effective environmental policies and regulations regarding pollution abatement. For example, for companies in industry sectors that have below–average return from EP, proper economic incentives can help them to be progressive in their EM activities (Clarkson et al., 2011). Also, for some specific industries with high a company–level variation regarding EP–FP relationship, voluntary instead of compulsory environmental programs and policies may be more flexible and effective for companies operating in those industries. Consulting service and suggestions can also be provided to companies with low EM capabilities and efficiencies to help them overcome resource or management constraints if there is any. Furthermore, since public recognition is important for companies with superior EP to achieve high–level FP, governments as well as other non–governmental organizations (NGOs) can help to provide incentives by increasing the public exposure of company's EP and improve the "green goodwill" of various stakeholders in the whole society.

7.2. Limitations and future research

The first limitation is about the data used in this study. Only large companies are included in this study without any medium or small ones. However, studies have shown that companies with different sizes may devote to environmental initiatives differently (Dierkes and Coppock, 1978; Trotman and Bradley, 1981), because benefits from EM activities are different for large and small companies (Dixon–Fowler et al., 2012). Large companies are more likely to adopt EM strategies in their decision making and benefit more from EM activities because they usually receive more attention from the public than small firms (Ashby et al., 2012; Hart and Ahuja, 1996; Pagell and Wu, 2009). Therefore, there could be some biases to a certain extent in analyzing the EP–FP relationship by using only large companies. That may explain why there is no variation regarding the impact of FP on EP across companies and industries. Therefore, one possible extension for future research is to include medium or small companies in the sample to test and verify the role of company size in determining the existence of heterogeneity regarding EP–FP relationship.

The second limitation is that this study only analyzes companies from the United States. The results found here might not be generalized to companies in other countries. Thereby another possible extension is to apply the multilevel framework to companies in other countries to investigate the EP–FP relationship. Moreover, location can be added to the multilevel framework as another important level because different locations (countries or states) may have different characteristics and requirements in their environmental policies and regulations (Jayaraman and Liu, 2019). However, adding locations into the multilevel framework will change the current nested structure into a cross–classified multilevel model, which will make the multilevel framework more complicated. For instance, one specific location (i.e. country or state) may include several industry sectors, while one industry sector can reside in several locations. This cross–classified multilevel modeling could be even more complicated if the interaction effect between industry and location is specified. For example, the EP–FP relationship in a specific industry in one location could be different from that in another location for the same industry. This industry–location interaction is vital for managers when considering business expansion domestically or internationally, because different EM strategies may be required for their branches in new locations.

The third limitation is that this study doesn't fully address the effect of environmental regulation and other external uncertainties. Environmental regulations are normally viewed as an instrument to increase firms' competitiveness by forcing firms to seek more efficient and sustainable operations with lower cost as well as green innovations, which can lead to enhanced corporate image, new market opportunities and higher profitability (Porter and van der Linde, 1995). External uncertainties, such as technology and market uncertainties, indicate various risks and unclear expectation of return regarding firms' investment in improving environmental performance (Chang, 2016; Li et al., 2019; Liu et al., 2017). For example, demand uncertainty of green product or lack of information regarding the changes of green technologies may hinder firms from devoting great efforts in improving EM and green innovations (Lopez-Gamero et al., 2011; Liu et al., 2018). Future research can include those omitted company- and industry-level features if possible and study the roles of these external factors in impacting the EP-FP relationship.

Appendix 1. List of Industry Sectors in the Dataset

211 Oil and Gas Extraction
212 Mining (except Oil and Gas)
213 Support Activities for Mining
221 Utilities
311 Food Manufacturing
312 Beverage and Tobacco Product Manufacturing
315 Apparel Manufacturing
322 Paper Manufacturing
324 Petroleum and Coal Products Manufacturing
325 Chemical Manufacturing
326 Plastics and Rubber Products Manufacturing
332 Fabricated Metal Product Manufacturing
333 Machinery Manufacturing
334 Computer and Electronic Product Manufacturing
335 Electrical Equipment, Appliance, Component Manufacturing
336 Transportation Equipment Manufacturing
339 Miscellaneous Manufacturing
423 Merchant Wholesalers, Durable Goods
424 Merchant Wholesalers, Nondurable Goods
441 Motor Vehicle and Parts Dealers

444 Building Material and Garden Equipment and Supplies Dealers
446 Health and Personal Care Stores
448 Clothing and Clothing Accessories Stores
452 General Merchandise Stores
481 Air Transportation
511 Publishing Industries (except Internet)
515 Broadcasting (except Internet)
517 Telecommunications
519 Other Information Services
522 Credit Intermediation and Related Activities
523 Securities, Commodity Contracts, Other Financial Activities
524 Insurance Carriers and Related Activities
531 Real Estate
541 Professional, Scientific, and Technical Services
561 Administrative and Support Services
621 Ambulatory Health Care Services
721 Accommodation
722 Services and Drinking Places

Appendix 2. Industry–Level Random EP Effects (p_j) on FP

Industry	effect	Standard error	rank
221	–.0001225	.000045	1
332	–.0001135	.000062	2
311	–.0000953	.000049	3
322	–.000091	.0000537	4
621	–.0000888	.000066	5
212	–.0000548	.0000618	6
531	–.0000521	.0000689	7
339	–.0000498	.0000684	8
312	–.0000413	.0000562	9
515	–.0000404	.0000582	10
561	–.0000338	.000073	11
326	–.0000328	.0000704	12
325	–.0000294	.0000362	13
315	–.000027	.000069	14
541	–.0000246	.0000574	15
335	–.0000105	.0000631	16
424	–.0000105	.0000725	17
448	−1.50e–06	.0000639	18
524	1.12e–06	.0000515	19
441	6.95e–06	.0000863	20
517	.0000101	.0000641	21
722	.0000113	.0000703	22
423	.0000117	.0000753	23
213	.0000121	.0000623	24
511	.0000132	.0000597	25
452	.0000188	.0000549	26
334	.0000222	.0000354	27
336	.0000224	.0000514	28
721	.0000239	.0000642	29
211	.0000266	.0000663	30
444	.0000572	.0000761	31
333	.0000598	.0000517	32
519	.000061	.0000758	33
523	.0000629	.0000641	34
481	.000071	.0000538	35
446	.0000793	.0000663	36
522	.0001686	.0000471	37
324	.0001794	.0000605	38

Appendix 3. Company–Level Random EP Effects (s_{ij}) on FP

company	effect	standard error	rank
027786	–.0002072	.0000863	1
002884	–.0001672	.0000777	2
003231	–.0001583	.0000875	3
003532	–.0001555	.000079	4
008030	–.0001504	.0000956	5
135990	–.0001304	.0000793	6

012726	–.0001278	.0000709	7
002154	–.0001216	.0000746	8
118577	–.0001151	.0000769	9
150937	–.0001051	.0000876	10
The rest 441 companies are here			
Available on request			
002817	.0001211	.0000914	452
006266	.0001211	.0000843	453
002136	.0001251	.0000974	454
008762	.0001253	.0000911	455
008530	.0001367	.0000934	456
011259	.0001433	.0000952	457
008479	.000152	.000092	458
005047	.0001629	.0000988	459
002968	.0001837	.0000909	460
001690	.0002135	.0000898	461

References

Aras, G., Aybars, A., Kuthu, O., 2010. Managing corporate performance: investigating the relationship between corporate social responsibility and financial performance in emerging markets. Int. J. Product. Perform. Manag. 59 (3), 229–254.

Ashby, A., Leat, M., Hudson–Smith, M., 2012. Making connections: a review of supply chain management and sustainability literature. Supply Chain Manag.: Int. J. 17 (5), 497–516.

Barbera, A., McConnell, V., 1990. The impact of environmental regulations on industry productivity: direct and indirect effects. J. Environ. Econ. Manag. 18 (1), 50–65.

Barney, J.B., Ouchi, W.G., 1986. Organizational Econom- Ics. Jossey-Bass, San Francisco.

Beatty, Timothy, Shimshack, Jay P., 2010. The impact of climate change information: new evidence from the stock market. Journal of Economic Analysis & Policy: Contributions 10 (1), 60.

Browne, W.J., Goldstein, H., Rasbash, J., 2001. Multiple membership multiple classification (MMMC) models. Stat. Model. 1 (2), 103–124.

Cahan, S.F., Chen, C., Chen, L., Nguyen, N.H., 2015. Corporate social responsibility and media coverage. J. Bank. Financ. 59, 409–422.

Chang, C.H., 2016. The determinants of green product innovation performance. Corp. Soc. Responsib. Environ. Manag. 23 (2), 65–76.

Christmann, P., 2000. Effects of "best practices" of environmental management on cost competitiveness: the role of complementary assets. Acad. Manag. J. 43 (4), 663–680.

Clarkson, M.B., 1995. A stakeholder framework for analyzing and evaluating corporate social Performance. Acad. Manag. Rev. 20 (1), 92–117.

Clarkson, P.M., Li, Y., Richardson, G.D., Vasvari, F.P., 2011. Does it really pay to be green? Determinants and consequences of proactive environmental strategies. J. Account. Public Policy 30 (2), 122–144.

Cohen, M., Fenn, S., Naimon, J., 1995. Environmental and Financial Performance: Are They Related? *Working Paper*. Vanderbilt University.

Cole, D.A., Maxwell, S.E., 2003. Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. J. Abnorm. Psychol. 112 (4), 558–577.

Curran, P.J., 2000a. A latent curve framework for the study of developmental trajectories in adolescent substance use. In: Rose, J.S., Chassin, L., Presson, C.C., Sherman, S.J. (Eds.), Multivariate Applications in Substance Use Research. Lawrence Erlbaum, Mahwah, NJ.

Datta, D., Guthrie, J., Wright, P., 2005. Human resource management and labor productivity: does industry matter? Acad. Manag. J. 48 (1), 135–145.

Derwall, J., Gunster, N., Bauer, R., Koedijk, K., 2005. The eco–efficiency premium puzzle. Financ. Anal. J. 61 (2), 51–63.

Dixon–Fowler, H.R., Slater, D.J., Johnson, J.L., Ellstrand, A.E., Romi, A.M., 2012. Beyond "Does it pay to be green?" A meta–analysis of moderators of the CEP–CFP relationship. J. Bus. Ethics 112 (2), 353–366.

Dierkes, M., Coppock, R., 1978. Europe tries the corporate social report. *Business and Society Review* Spring 21–24.

Donaldson, L., 2001. The Contingency Theory of Organizations. Sage, Thousand Oaks.

Eisenhardt, K.M., Martin, J.A., 2000. Dynamic capabilities: what are they? Strateg. Manag. J. 21 (10–11) 1105 1121.

Elsayed, K., Paton, D., 2005. The impact of environmental performance on firm performance: static and dynamic panel data evidence. Struct. Chang. Econ. Dyn. 16, 395–412.

Fatemi, A., Glaum, M., Kaiser, S., 2017. ESG performance and firm value: the moderating role of disclosure. Glob Financ. J. 38, 45–64.

Figge, F., Hahn, T., Schaltegger, S., Wagner, M., 2002. The sustainability balanced scorecard –linking sustainability to business strategy. Bus. Strateg. Environ. 11 (5), 269–284.

Fitzmaurice, G.M., Laird, N.M., Ware, J.H., 2012. In: Applied Longitudinal Analysis. John Wiley & Sons, Hoboken, NJ.

Garcia–Castro, R., Ariño, M.A., Canela, M.A., 2010. Does social performance really lead to financial performance? Accounting for endogeneity. J. Bus. Ethics 92 (1), 107–126.

Ghoul, S.E., Guedhami, O., Kowk, C.C.Y., Mishra, D.R., 2011. Does corporate social responsibility affect the cost of capital? J. Bank. Financ. 35 (9), 2388–2406.

Goss, A., Roberts, G.S., 2011. The impact of corporate social responsibility on the cost of bank loans. J. Bank. Financ. 35 (7), 1794–1810.

Griffin, J.J., Mahon, J.F., 1997. The corporate social performance and corporate financial performance debate: twenty–five years of incomparable research. Bus. Soc. 36 (1), 5–31.

Gupta, S., Goldar, B., 2005. Do stock markets penalize environment–unfriendly behavior? Evidence from India. Ecol. Econ. 52, 81–95.

Hansen, G.S., Wernerfelt, B., 1989. Determinants of firm performance: the relative importance of economic and organizational factors. Strateg. Manag. J. 10 (5), 399–411.

Hansen, D.R., Mowen, M.M., Guan, L., 2007. Cost Management: Accounting and Control, sixth ed. South–Western College Publishing, Mason, OH.

Hart, S.L., 1995. A natural–resource–based view of the firm. Acad. Manag. Rev. 20 (4), 996–1014.

Hart, S.L., Ahuja, G., 1996. Does it pay to be green? An empirical examination of the relationship between emission reduction and firm performance. Bus. Strateg. Environ. 5 (1), 30–37.

Hedström, Peter, Ylikoski, Petri, 2010. Causal mechanisms in the social sciences. Annu. Rev. Sociol. 36, 49–67.

Hoffman, L., 2015. Longitudinal Analysis: Modeling Within–Person Fluctuation and Change. Routledge Academic, New York, NY.

Horváthová, E., 2010. Does environmental performance affect financial performance? A meta-analysis. Ecol. Econ. 70 (1), 52–59.

Hou, K., Robinson, D.T., 2006. Industry concentration and average stock returns. J. Financ. 61 (4), 1927–1956.

Hull, C.E., Rothenberg, S., 2008. Firm performance: the interactions of corporate social performance with innovation and industry differentiation. Strateg. Manag. J. 29 (7), 781–789.

Jacobs, B.W., Singhal, V.R., Subramanian, R., 2010. An empirical investigation of environmental performance and the market value of the firm. J. Oper. Manag. 28 (5), 430–441.

Jayaraman, V., Liu, Z., 2019. Aligning governance mechanisms with task features to improve service capabilities—an empirical study of professional service outsourcing in India. Operations Management Research 12 (1–2), 19–39.

Judge, W., Douglas, T., 1998. Performance implications of incorporating natural environmental issues into the strategic planning process: an empirical assessment. J. Manag. Stud. 35 (2), 241–262.

King, A., Lenox, M., 2001. Lean and green? An empirical examination of the relationship between lean production and environmental performance. Prod. Oper. Manag. 10 (3), 244–256.

Klassen, R.D., McLaughlin, C.P., 1996. The impact of environmental management on firm performance. Manag. Sci. 42 (8), 1199–1214.

Klassen, R.D., Whybark, D.C., 1999. The impact of environmental technologies on manufacturing performance. Acad. Manag. J. 42 (6), 599–615.

Kogut, B., Zander, U., 1992. Knowledge of the firm, combinative capabilities and the replication of technology. Organ. Sci. 3 (3), 383–397.

Konar, S., Cohen, M.A., 2001. Does the market value environmental performance? Rev. Econ. Stat. 83 (2), 281–289.

KPMG, 2011. International survey of corporate responsibility reporting 2011. http://www.kpmg.com/Global/en/IssuesAndInsights/ArticlesPublications/corporate-responsibility/Documents/2011-survey.pdf Accessed on May 30th, 2016.

Lengnick–Hall, C.A., Wolff, J.A., 1999. Similarities and contradictions in the core logic of three strategy research streams. Strateg. Manag. J. 20 (12), 1109–1132.

Li, P., Menon, M., Liu, Z., 2019. Green innovation under uncertainty – a dynamic perspective. Int. J. Serv. Econ. Manag. 10 (1), 68–88.

Liu, Z., Jayaraman, V., Luo, Y., 2017. The unbalanced indirect effects of task characteristics on performance in professional service outsourcing. International Journal of Production Economics 193, 281–293.

Liu, Z., Wang, H., Li, P., 2018. The antecedents of green information system and impact on environmental performance. Int. J. Serv. Econ. Manag. 9 (2), 111–124.

Lopez-Gamero, M.D., Molina-Azorin, J., Claver-Cortes, E., 2011. Environmental uncertainty and environmental management perception: a multiple case study. J. Bus. Res. 64 (4), 427–435.

McGuire, J.B., Schneeweiss, T., Branch, B., 1990. Perceptions of firm quality: a cause or result of firm performance. J. Manag. 16 (1), 167–180.

Melville, N., Gurbaxani, V., Kraemer, K., 2007. The productivity impact of information technology across competitive regimes: the role of industry concentration and dynamism. Decis. Support Syst. 43 (1), 229–242.

Moneva, J.M., Cuellar, B., 2009. The value relevance of financial and non–financial environmental reporting. Environ. Resour. Econ. 44, 441–456.
Moore, G., 2001. Corporate social and financial performance: an investigation of the UK supermarket industry. J. Bus. Ethics 34, 299–315.
Nelling, E., Webb, E., 2009. Corporate social responsibility and financial performance: a virtuous circle revisited. Rev. Quant. Financ. Account. 32 (2), 197–209.
Nelson, R.R., Winter, S.G., 1982. An Evolutionary Theory of Economic Change. Belknap Press, Cambridge, MA.
Ngwakwe, C.C., 2009. Environmental responsibility and firm performance: evidence from Nigeria. Int. J. Humanit. Soc. Sci. 3, 97–104.
Orlitzky, M., Schmidt, F.L., Rynes, S.L., 2003. Corporate social and financial performance: a meta–analysis. Organ. Stud. 24 (3), 403–441.
Pagell, M., Wu, Z., 2009. Building a more complete theory of sustainable supply chain management using case studies of 10 exemplars. J. Supply Chain Manag. 45 (2), 37–56.
Patterson, B.F., 2013. A cross–classified multilevel model for first–year college natural science performance using SAS. In: Garson, G.D. (Ed.), Hierarchical Linear Modeling: Guide and Applications. Sage Publications, Thousand Oaks, CA, pp. 291–310.
Pfeffer, J., 1997. New Directions for Organization Theory. Oxford University Press, New York, NY.
Porter, M., 1991. America's green strategy. Sci. Am. 264 (4), 168.
Porter, M., van der Linde, C., 1995. Toward a new conception of the environmental competitiveness relationship. J. Econ. Perspect. 9 (4), 97–118.
Priem, R.L., Butler, J.E., 2001. Is the resource–based ‘view’ a useful perspective for strategic management research? Acad. Manag. Rev. 26 (1), 22–40.
Raudenbush, S.W., Bryk, A.S., 2002. Hierarchical Linear Models: Applications and Data Analysis Methods ume 1 Sage Publications, Thousand Oaks, CA.
O'Reilly, C.A., Tushman, M.L., 2008. “Ambidexterity as a dynamic capability: resolving the innovator's dilemma. Res. Organ. Behav. 28, 185–206.
Russo, M.V., Fouts, P.A., 1997. A resource–based perspective on corporate environmental performance and profitability. Acad. Manag. J. 40 (3), 534–559.
Singer, J.D., Willett, J.B., 2003. In: Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. Oxford university press, New York.
Stanwick, P., Stanwick, S., 1998. The relationship between corporate social performance, and organizational size, financial performance, and environmental performance: an empirical examination. J. Bus. Ethics 17 (2), 195–204.
Teece, D.J., 1984. Economic analysis and strategic management. Calif. Manag. Rev. 26 (3), 87–110.
Teece, D., Pisano, G., Shuen, A., 1997. Dynamic capabilities and strategic management. Strateg. Manag. J. 18 (7), 509–533.
Trotman, K., Bradley, G., 1981. Associations between social responsibility disclosure and characteristics of companies. Account. Org. Soc. 6, 355–362.
Ullmann, A., 1985. Data in search of a theory: a critical examination of the relationships among social performance, social disclosure, and economic performance of US firms. Acad. Manag. Rev. 10 (3), 540–557.
Waddock, S.A., Graves, S.M., 1997. The corporate social performance–financial performance link. Strateg. Manag. J. 18 (4), 303–319.
Wagner, M., Schaltegger, S., Wehrmeyer, W., 2001. The relationship between the environmental and economic performance of firms: what does theory propose and what does empirical evidence tell us? Greener Manag. Int. 34, 95–108.
Walley, N., Whitehead, B., 1994. It's not easy being green. Harv. Bus. Rev. 72, 46–52.
Wernerfelt, B., 1984. A resource–based view of the firm. Strateg. Manag. J. 5, 171–180.
Wood, D.J., Jones, R.E., 1995. Stakeholder mismatching: a theoretical problem in empirical research on corporate social performance. Int. J. Organ. Anal. 3 (3), 229–267.
Zott, C., 2003. Dynamic capabilities and the emergence of intraindustry differential firm performance: insights from a simulation study. Strateg. Manag. J. 24 (2), 97–112.
https://www.davispolk.com/files/2017-07-12_esg_reports_ratings_what_they_are_why_they_matter_0.pdf, Accessed date: 16 November 2017.