PRI Academic Conference 2010
Mainstreaming Responsible Investment
Copenhagen Business School
May 5-7, 2010

Corporate Environmental Performance: Consistency of Metrics and Identification of Drivers

Natalia Semenova
Åbo Akademi University and
the Graduate School of Accounting at the Academy of Finland

Address for correspondence: Natalia Semenova, School of Business and Economics, Åbo Akademi University, FI-20500 Turku. Email: natalia.semenova@abo.fi

March 2010

Corporate Environmental Performance:
Consistency of Metrics and Identification of Drivers

Abstract

In order to meet the increased demand for environmental information professional rating services, and regulatory agencies have developed a range of environmental metrics that differs from simple narrow indicators to comprehensive measurement systems. This study is among the first to provide insight into the assessment of the construct validity of prominent environmental performance ratings. Using a set of environmental dimensions in KLD, GES, and ASSET4 ratings, this study indicates that aggregate environmental performance metrics sufficiently correlate and provide consistent information when comparing companies. The KLD environmental rating *concerns* is particularly likely to provide good summaries of the environmental impact of industrial activities in contrast to the KLD environmental performance measure *strengths*. The observed different pattern in KLD environmental dimensions suggests that they are distinct construct and should not be combined in future research. This paper defines environmental industry risk, impact or concerns as being factors that drive corporate environmental performance. The contribution of this paper is, therefore, a validation of environmental ratings and a sharper focus upon factors that are associated with high levels of environmental performance. The findings for *MSCI U.S. World companies 2003-2008* have implications for measurement theory and stakeholder theory. In addition, this study discusses the implications of findings for advocates and sceptics of environmental ratings, as well as for academics and practitioners in the realm of SRI and CSR.

Keywords: information disclosure; environmental performance; ratings; convergent validity

Preliminary draft: please do not quote.

This version: March 2010

1. Introduction

An important issue that faces many publicly traded companies today is the disclosure of relevant environmental information to assist stakeholders in making informed decisions. To date, a rapid growth of socially responsible investors requires related information that identifies companies deemed acceptable for investments; this detailed information comes in form of environmental, social, and corporate governance performance. To a certain extent, accounting guidelines on environmental disclosures direct companies to disclose environmental liabilities as risks associated with the impact of environmental legislation and regulations. The existing literature on the value relevance of environmental liabilities found that such environmental performance information is relevant to investors. However, another stream of research that explores the relationship between environmental disclosures and environmental performance is concerned with the lack of quality in environmental performance disclosures in annual reports, 10K forms, and sustainability/environmental reports (Barth and McNichols, 1994; Clarkson et al., 2008). The environmental disclosures that the Securities and Exchange Commission and the Financial Accounting Standard Board require have provided little direction (Gamble et al., 1995) and have been largely discretionary (Blacconiere and Northcut, 1997). Disclosure of environmental performance indicators based upon the Global Reporting Initiative sustainability reporting guidelines has only limited usefulness in measuring the environmental performance of a given company (Barth and McNichols, 1994; Clarkson et al., 2008; Schäfer, 2005). As a communication tool, environmental disclosures enhance transparency by bringing a positive environmental profile to companies and strengthening their relations with stakeholders (Azzone et al., 1996; Blacconiere and Northcut, 1997). Since extensive environmental disclosures are more likely to be interpreted by investors as a positive sign that a company manages its environmental activities (Blacconiere and Northcut, 1997), a crucial component of these disclosures is that unverifiable practices of successful companies can be manipulated and misinterpreted or easily mimicked by environmental performers. The dispersion that is present in corporate environmental disclosures and a lack of comparable standards do not facilitate effective assessment and comparison of environmental performance across companies (Ilinitch et al., 1998; Bassen and Kovacs, 2008). In a situation such as this, different environmental performance metrics provided by professional services serve the information needs of users.

The difficulty in assessing environmental performance has been widely documented in the literature (Azzone et al., 1996; Ilinitch et al., 1998; Schäfer, 2005). Corporate environmental performance as a complex construct aggregates multiple types of indicators and can be split into different measurement areas (Ilinitch et al., 1998; William and Welford, 1998; Azzone et al., 1996). The assessment of companies on environmental issues has evolved into the establishment of special professional research services (ratings) that develop specific evaluation methods to rate and rank companies; examples of such research services or ratings are KLD, Innovest, ASSET4, GES, and Vigeo (Schäfer et al., 2006). The aim of research services is to overcome the information asymmetry between companies and their stakeholders/investors by disclosing performance information aiming at a high level of reliability and validity (Chatterji and Toffel, 2010; Waddock, 2003). A general structural model of environmental, social, and corporate governance ratings developed by Schäfer (2005) brings to light a process of screening and monitoring company stocks, as conducted by the rating institutions for capital markets. The measuring methods and assessment procedures in this process are based upon international standards, conventions, and paradigms of corporate responsibility. For a variety of reasons, academics and stakeholders questioned whether such environmental metrics that are proliferated in the absence of clear and generally accepted guidelines convey an accurate assessment of corporate environmental performance (Ilinitch and Schaltegger, 1995; Ilinitch et al., 1998; Entine, 2003; Chatterji and Levine, 2006). Moreover, it is becoming more difficult for users to differentiate between the increasing number of environmental metrics and standards (although a certain degree of consolidation of professional services has more recently taken place).

This paper addresses the concern in the academic literature regarding a systematic and empirical validation of environmental metrics (Chatterji and Levine, 2006; Chatterji et al., 2009). By using a rigorous research design, the purpose is to present a convergent validity study comparing commonly accepted KLD, GES, and ASSET4 environmental metrics. In order to accomplish this objective, this study first correlates different combinations of environmental ratings to test the extent to which they converge to a uniform metric. This analysis enables the differentiation between the results of two constructs: namely the environmental impacts of industrial activities and the company's environmental actions. Additionally, this study looks beyond average effects and identifies the consistency of environmental metrics across industries. A final focus explores industry risk as a factor that drives company decisions to invest in environmental performance. Polluting industries are more regulated than clean industries and have different environmental constraints upon company operations that can directly affect the incentives for corporate environmental performance. Using a stakeholder framework, this study relates the industry-specific environmental risk and company-specific characteristics (such as size and profitability) to corporate environmental performance.

In brief, the findings of the study indicate that KLD environmental strengths, GES, and ASSET4 environmental performance metrics are consistent aggregate metrics and converge on the same construct among MSCI U.S. companies (2003-2008). In contrast, KLD environmental concerns tend to be a risk-oriented metric of environmental activities that closely relates to GES environmental industry risk. Environmental performance in KLD, GES, and ASSET4 metrics provides consistent information to compare companies across industries. Industry-specific risk and company-specific size factors are associated with high levels of corporate environmental performance. Consequently, this paper is among the first to consider the convergent validity of prominent environmental ratings and to show that industry risk is an important factor in determining corporate environmental performance.

This research provides a platform for evaluation of environmental metrics for stakeholders when rewarding or punishing companies on the basis of environmental performance. Stakeholders and investors who rely upon environmental ratings to identify target companies might be mislead if the prevailing environmental ratings do not provide an accurate assessment of corporate environmental performance. The differences in corporate environmental performance ratings can confuse stakeholders that are attempting to use certain dimensions of the data in order to compare performance across companies and industries. This work also provides a platform for the substantial theoretical and empirical literature that utilizes environmental performance measurement theory. More than 100 studies have examined the link between corporate social/environmental performance and financial performance (Margolis et al., 2007). Identifying the empirical divergence between KLD environmental strengths and concerns indicates that some conclusions from prior empirical research could require reinterpretation. Evaluation of environmental metrics allows for a better understanding of the relationships between environmental drivers and outcomes, and hopefully leads to more consistent interpretation of research results.

The paper is organized as follows: the next section provides a literature review and hypotheses development, followed by a description of methodology of empirical tests and a presentation of environmental metrics. The paper then presents the results of univariate and multivariate tests. The final section summarizes the study’s main findings complete with a discussion of the practical implications and suggestions for future research.

2. Literature Review and Hypothesis

2.1 Literature Review

The existing literature on corporate environmental performance can be categorized into three broad groups. The first group of research develops theoretical frameworks, specifying the types of information and common dimensions/factors underlying environmental performance constructs (Azzone et al., 1996; Tyteca, 1996; Wood, 1991; Lober, 1996; Ilinitch et al., 1998). This scope of these studies indicates that the concept of environmental performance is a global and aggregate theoretical construct that provides a summary measurement by aggregating two perspectives of the environment: the effect of industrial activities (risk or impact factor) and the environmental actions taken on behalf of the company to control for company-specific environmental risk (environmental performance). The second line of literature incorporates environmental performance metrics into research methodology: e.g. the Toxic Release Inventory emissions indicators (King and Lenox, 2001; Konar and Cohen, 2001; Clarkson et al., 2008); the KLD index of social/environmental performance (Hillman and Keim, 2001; Statman and Glushkov, 2009; Baron et al., 2009); and, the Innovest eco-efficiency ratings (Derwall et al., 2005; Guenster et al., 2006). This group of research finds that such environmental performance information is valuable to investors seeking extra-financial returns and stakeholders concerned with environmental conflicts. A third line of thought, which is most relevant to this study, explores the validity of corporate social/environmental metrics.

Through a comparison of data from the *Fortune* reputation survey and the holdings list of the best known “social choice” mutual funds, Sharfman (1996) examined a construct validity of KLD corporate social performance single metric. The study was concerned with a lack of specific theory that was used to develop KLD criteria. The authors correlated different combinations of the KLD ratings with three sets of measures of overall corporate social performance. The results of correlation analysis indicate that KLD social performance ratings have a moderate level of correlations with the other measures of corporate social performance, thereby, measuring at least part of the same construct.

By applying the explanatory factor analysis procedures, Mattingly and Berman (2006) conducted a methodological study that classifies KLD corporate social performance dimensions into four latent constructs derived from the extant typologies of corporate social actions. In taxonomy of KLD ratings, the authors made a distinction between social actions toward technical stakeholders and social actions toward institutional stakeholders. The concurrent validity of the resulting classifications was then examined based upon the correlations with variables known to be associated with corporate social performance. The empirical results indicated that environmental weakness, community weakness, and environmental strength related to the latent factor of institutional weakness. The latent factor of institutional strength consisted of KLD diversity and community strength dimensions. The third latent factor, that of technical weakness, was associated with employee, diversity, product, and governance weakness dimensions. Technical strength included loadings from product, governance, and employee strengths. An especially important observation for this study is that KLD environmental strength and weakness converge in the same latent factor (such as institutional strength) and are highly correlated. Taken together, a key finding of the study by Mattingly and Berman is that KLD weakness and strength dimensions are both empirically and conceptually distinct constructs of corporate social action.

Chatterji et al. (2009) explored the validity of KLD environmental ratings in terms of providing transparency about past and future environmental performance. In order to assess the extent to which KLD environmental ratings rendered transparency of past performance, the authors regressed three disaggregated KLD concerns ratings (such as Hazardous waste, Regulatory problems, and Substantial emissions) and total environmental concerns score on several environmental performance metrics, namely: emissions, penalty values, number of violations, number of major spills, number of permit denials or shut-ins, and a dummy that indicates whether each performance metric is zero or has a positive value. The regression results indicated that KLD environmental concerns metric provides information about past corporate environmental performance. KLD measures were found to be correct in identifying companies that have high emissions and regulatory concerns when not normalizing by size, while they were only modestly effective when size adjustments were performed. Chatterji et al. (2009) developed a model that predicts environmental performance metrics based on KLD environmental ratings in order to assess the extent to which KLD environmental ratings provided transparency about future environmental performance. Evidence showed that the KLD single net environmental score is significantly negative in predicting pollution levels, the value and number of regulatory penalties, and number of major spills. However, KLD environmental concern metric drives these relationships. No evidence was found to support the notion that KLD environmental strengths predict subsequent environmental outcomes.

Ilinitch et al. (1998) developed dimensions of the environmental performance construct and investigated reliability and validity of environmental performance indicators across different data sources. Principle component analysis was used to specify indicators that related to four dimensions of the environmental performance construct: internal systems, external stakeholder relations, external impacts, and internal compliance. The primary source of analysis was the Investor Responsibility Research Center Corporate Environmental Profiles (IRRC) reports. Environmental performance indicators were as follows: absolute and size-weighted CEP's total TRI emissions, IRRC's Efficiency Index defined as a sized-weighted total emissions variable, a penalty based variable IRRC's Compliance Index, and IRRC's Spill Index[1]. The binary variable

[1] TRI is Toxic Release Inventory and CEP is the Council on Economic Priorities

of organizational process was constructed based upon measurements where the company used environmental TQM measures, had written environmental policies, conducted environmental audits, compensated CEOs for environmental performance, and applied US standards to overseas operations. All measures were converted into a five-point ranking scheme. The study demonstrated the variability and inconsistency across different environmental performance indicators by using a sample of eleven large petroleum companies. The inconsistency among indicators was due to the differences in denominators of emission variables, the assessment of scores with or without an industry factor, and the lack of variance in measures such as spills, compliance, and organizational process. In the following analysis, composite environmental performance ratings were used (such as Franklin Research and Development Corporation's rating and CEP's Shopping for a Better World rating). The regression analysis of environmental performance ratings indicates that each environmental rating is related to different aspects of environmental performance (such as regulatory compliance and environmental impacts); additionally, a public view factor influences both measures.

In summary, limited and low validity of environmental metrics has been given in prior research. Most of the studies focused merely upon the validation of KLD metrics. The environmental benchmarks that have been used raise serious concerns about their subjective assessment, the dominance of financial factor, the narrow context, and the limitations by certain industries (Tyteca, 1996; Fryxell and Wang, 1994; Brown and Perry, 1994; King and Lenox, 2001). It still remains unclear whether environmental performance ratings are actually consistent and help stakeholders in identifying environmentally responsible companies. Prior research suggests that further work is needed to better understand the environmental metrics that are already available (Mattingly and Berman, 2006; Ilinitch et al., 1998; Chatterji et al., 2009; Sharfman, 1996). The analysis performed in this study is among the first to empirically examine the extent to which well known multi-dimensional KLD, GES, and ASSET4 environmental ratings converge. The ratings are widely used by both academics and investors in the realm of SRI and CSR.

2.2 Hypothesis Development

Given the fact that any measure reflects an underlying theoretical concept (Bagozzi et al., 1991), this study aims to develop a hypothesis of consistency among environmental performance metrics. Numerous measures and scales have been used to operationalise the corporate environmental performance construct (see Margolis et al., 2007 for a review). One of the most frequently used measures of environmental performance - especially in the US context - is the pollution emission from the Toxic Release Inventory and the Council on Economic Priorities database (or alternatively, the ratio of pollution emissions to total company sales/revenues). The pollution emission is an outcome metric used to measure a company's waste generation and pollution reduction activities that is required to be consistently disclosed across companies. Studies utilised the pollution emission metric for a number of purposes (e.g. (1) to investigate stock market reaction to the announcement of the pollution news (Hamilton, 1995; Khanna et al., 1998); (2) to explore the relationship between emissions reduction and economic/financial performance (Hart and Ahuja, 1996; King and Lenox, 2002; King and Lenox, 2001; Konar and Cohen, 2001; Clarkson et al., 2006); (3) to examine the associations between environmental performance and environmental disclosure (Clarkson et al, 2008; Patten, 2002); and (4) to measure environmental efficiency or *eco-efficiency* of a company (Chatterji and Toffel, 2010). Although the pollution emission is output oriented and officially sanctioned, the metric is limited in many ways. Gerde and Logsdon (2001) and Ziegler et al. (2007) raise the general concern that emissions data are not appropriate proxies for the general phenomenon of environmental performance since the said data relate to only one type of environmental concern. Elsayed and Paton (2005) and Stanwick and Stanwick (1998) criticize the pollution emission environmental measure because it does not allow investigating environmental performance of companies in low polluting industries. King and Lenox (2001) and Ziegler et al. (2007) argue that the pollution emission metric confounds pollution that results from industry activities with pollution that results from poor environmental management. This research relies upon common multi-dimensional metrics of corporate environmental performance since environmental performance is conceptually considered to be a complex aggregate construct (Ilinitch et al., 1998) and a stream of research that uses pollution emission metric is limited by certain industries.

KLD environmental metrics have been extensively used in scholarly research to operationalise the environmental performance construct. Leading scholars have called the KLD data as the research standard for measuring corporate social performance (Waddock, 2003; Mattingly and Berman, 2006; Chatterji and Toffel, 2010). KLD uses positive and negative screens to monitor corporate environmental performance. Positive screens indicate environmental strengths of a company; negative screens indicate environmental weaknesses or concerns of the company. Each screen is summarized in a binary variable, which reflects whether the company meets the particular criterion or not. To arrive at a single net environmental score, the common practice used in academic literature is to subtract the concerns from the strengths (Chatterji et al., 2009; Statman and Glushkov, 2009; Waddock and Graves, 1997; Hillman and Keim, 2001; Griffin and Mahon, 1997). Additionally, studies analyzed KLD environmental metrics separately. The KLD environmental concerns dimension has been applied to differentiate better and worse environmental performers (Cho and Patten, 2007) in order to indicate companies with poor environmental performance ratings (Chatterji and Toffel, 2010). Companies receiving the KLD environmental concerns rating are classified as worse environmental performers or as poorly rated companies; companies with no KLD environmental concerns ratings are identified as having better environmental performance or as well rated companies. Taken together, KLD environmental concerns are labelled as a measure of poor environmental performance. KLD environmental metrics have been used in a variety of the research settings: e.g. (1) to investigate the relationship between corporate social/environmental and financial performance (Scholtens and Zhou, 2008; Berman et al., 1999); (2) to explore the effect of socially responsible investing on portfolio performance (Kempf and Osthoff, 2007; Statman and Glushkov, 2009); (3) to compare the ratings with reputations and their attractiveness as employees (Turban and Greening, 1997); (4) to test the effects of institutional investor types and governance upon corporate social/environmental performance (Johnston and Greening, 1999); and (5) to examine how companies adjust their environmental performance in response to ratings (Chatterji and Toffel, 2010). Given its breadth, the KLD database holds potential for being a commonly accepted set of environmental performance metrics.

More recently, researchers have continued to develop the field of corporate environmental performance by utilizing other environmental metrics. Multi-dimensional Innovest eco-efficiency metric (Derwall et al., 2005, Guenster et al., 2006) operationalised the concept of *eco-efficiency*, which is defined as maximizing the effectiveness of business processes while minimizing their impact on the environment (Sinkin et al., 2008). Derwall et al. (2005) found that the portfolio of high-ranked eco-efficient companies provide higher risk-adjusted average returns than its low-ranked counterpart. Although the results seem to contradict the risk-return paradigm, Statman and Glushkov's (2009) study shows that the portfolio of companies with best-in-class KLD environmental scores outperforms the portfolio of companies with worst-in-class scores. Semenova and Hassel (2008) and Olsson (2007) have developed a multi-dimensional view of environmental management by using GES company-specific environmental metrics. Semenova and Hassel (2008) found that the reputational benefits of environmental preparedness increased market value and environmental performance also brought operational benefits to financial performance. The analysis performed by Hillman and Keim (2001) and Berman et al. (1999) indicates that the KLD composite environmental dimension is not related to operating performance and market value added. Scholtens and Zhou (2008) found that KLD environmental concerns have a detrimental impact on stock (equity) return earned by stakeholders and the concerns increased financial risk (defined as volatility of the stock prices). They also determined that KLD environmental strengths do not relate to financial performance or financial risk. Olsson (2007) constructed high and low risk portfolios based upon combinations of GES industry-specific and company-specific environmental risks. The study indicates that neither high nor low risk portfolio produced abnormal returns. By applying negative and positive screening policies to the KLD universe, Kemp and Osthoff (2007) found that the long-short strategy yielded an insignificant alpha, based upon an environmental screen. This means that studies that have been conducted by utilizing different environmental performance metrics have provided consistent evidence on market returns from corporate environmental performance in different methodological settings. Therefore, it is expected that multi-dimensional environmental performance metrics will reflect the same environmental concept. As a result, the following hypothesis is proposed:

Hypothesis 1: Environmental performance metrics are highly related, ceteris paribus.

Much of the prior research has focused upon financial benefits as incentives for company managers to invest in corporate environmental performance and disclosure. A growing body of empirical studies demonstrates that better environmental performers could gain higher reputational benefits than worse environmental performers. More recently, Artiach et al. (2010) investigated the company-specific factors associated with corporate sustainability performance. The authors found consistent results that company size is associated with high levels of corporate sustainability performance and limited evidence that the level of profitability drives corporate social performance. In general, large companies have more resources and are more likely to exhibit good relations with stakeholders than small companies would (Waddock and Graves, 1997; Scholtens and Zhou, 2008). Clarkson et al. (2006) found that companies that improve environmental performance in the polluting industries, tend to have greater growth and higher Tobin's Q immediately prior to the decision to improve. Chatterji and Toffel (2010) argue that company management would respond to poor initial environmental ratings by taking action on improving environmental performance. In particular, those companies operating in industries with high environmental constrains that receive negative environmental ratings are more motivated to improve their subsequent environmental performance. Although this research explored the moderator effect of regulatory environmental performance upon corporate environmental performance, it still leaves unanswered the question of how industry context directly affects corporate environmental performance. By using univariate t-test, Semenova and Hassel (2008) provided some evidence that companies that operate in polluting industries have significantly higher environmental preparedness and performance than those that operate in green industries. According to Cho and Patten (2007), companies from environmentally sensitive industries have incentives for greater environmental disclosures than companies from non-environmentally sensitive industries. The following hypothesis is proposed, which is consistent with prior findings:

> *Hypothesis 2: Environmental industry risk is positively related to corporate environmental performance, ceteris paribus.*

3. Methodology and Data

3.1 Test of Hypothesis 1

To assess the validity of the environmental performance metrics, this study employs multiple environmental variables measured by different methods and multiple empirical tests[2]. The use of multiple environmental variables permits taking into account random and systematic errors in measurement methods and differentiating between underlying environmental constructs. Multiple empirical tests allow distinguishing between substantive variance of environmental measures and the variance of irrelevant factors when compiling measurements and unobservable factors, which measurement methods have omitted. The main environmental metrics meet the following criteria: first, they are composite environmental measures that aggregate various one-dimensional indicators; and secondly, the environmental metrics do not require any weight on the environmental indicators that are aggregated. Although these indicators may not be equivalent, deriving individual weights is not always advisable since it renders the interpretation of results meaningless and complex (Tyteca, 1996; Chatterji and Levine, 2006). As a third criterion, environmental metrics follow the negative screening policy and exclude controversial business areas: such alcohol, tobacco, gambling, military, firearms, and nuclear power (Kempf and Osthoff, 2007). A forth criterion is that the assessment of environmental metrics is done on objective and reliable information, as well as is less susceptible to self-reported biases.

Following Sharfman (1996), the construct validation aimed to assess how different measures represent its latent constructs (concept) has been used in this study[3]. The approach for assessing construct validity is known as convergent validity that is a subcategory of criterion-related

[2] Validity means "whether the measure identifies performance that is important to society" (Chatterji and Levine, 2006, p.33). "A measure is reliable if it provides the same answer when applied more than one time " (Chatterji and Levine, 2006, p.32)
[3] Bagozzi et al. (1991, p.421) defines construct validity as "the extent to which an operationalisation measures the concept it is supposed to measure".

validity. According to this type of validation, one observes that the measures of constructs that should be theoretically related to each other, in fact, relate to each other empirically. The correlation matrix is typically used to estimate the degree to which any two measures relate to each other. Convergent correlations between theoretically similar measures should always be higher than correlations between theoretically dissimilar ones. Overall, convergent validity provides evidence to conclude that different environmental metrics are related to the same construct.

Additionally, this paper looks beyond average effects and identifies the consistency of environmental metrics across different industries. The study conducts univariate comparisons of environmental variables for a group of companies in polluting industries with a control sample of companies in green industries. A t-test of means is used to statistically compare differences in the mean of environmental metrics across companies in polluting and green industries.

Exploring the environmental metrics in a multivariate setting complements the univariate tests since measurement errors and unobservable factors provide potential threat to the validity of research findings. This study proposes to estimate the following linear econometric models using panel data:

$$EINDEX_{i,t} = \beta_0 + \beta_1 ESTR_{i,t} + \beta_2 ECON_{i,t} + \beta_3 EPILAR_{i,t} + u_i + \varepsilon_{i,t} \tag{1}$$

$$EPILAR_{i,t} = \beta_0 + \beta_1 ESTR_{i,t} + \beta_2 ECON_{i,t} + \beta_3 EINDEX_{i,t} + u_i + \varepsilon_{i,t} \tag{2}$$

$$ESTR_{i,t} = \beta_0 + \beta_1 EINDEX_{i,t} + \beta_2 ECON_{i,t} + \beta_3 EPILAR_{i,t} + u_i + \varepsilon_{i,t} \tag{3}$$

where $EINDEX_{i,t}$ is GES environmental performance index; $ESTR_{i,t}$ is KLD environmental strengths; $ECON_{i,t}$ is KLD environmental concerns; $EPILLAR_{i,t}$ is ASSET4 environmental pillar. The value of i denotes a company and t indicates time periods. The term u_i captures random variables related to unobservable factors. The constant terms (β_1 and β_3) that correspond to the coefficients of environmental metrics in Equations (1) - (3) are expected to be positive. The coefficient of KLD environmental concerns (β_3) is expected to be negative. The KLD, GES, and ASSET4 environmental metrics are introduced below.

3.2 Test of Hypothesis 2

Hypothesis 2 is tested to gain insight into the factors that drive high levels of corporate environmental performance. The foundation for the empirical test is the regression of environmental performance on environmental industry risk, company size, and profitability.

$$EPERF_{i,t} = \beta_0 + \beta_1 IRISK_{i,t} + \beta_2 SIZE_{i,t} + \beta_3 PROF_{i,t} + u_i + \varepsilon_{i,t} \qquad (4)$$

where $EPERF_{i,t}$ is environmental performance. $IRISK_{i,t}$ is environmental industry risk; $SIZE_{i,t}$ is company size measured as the natural log of total assets; $PROF_{i,t}$ representing profitability is measured by two alternative proxies: such as return on assets (ROA) and return on equity (ROE)[4]. Profitability and size data have been obtained from Thomson Financial DataStream. The constant terms (β_1, β_2, and β_3) in Equation (4) are expected to be positive.

3.3 Environmental Metrics

Environmental metrics have been obtained from KLD Research & Analytics, Inc. (KLD), GES Investment Services Risk Rating (GES), and ASSET4: Thomson Reuters rating databases.

KLD ratings have been used extensively in studies on CSR and SRI (Chatterji and Toffel, 2010) and considered to be reliable, broad-ranging, and valid environmental performance constructs

[4] ROA is estimated as operating income divided by total assets. ROE is calculated as operating income divided by common shareholder's equity

(Sharfman, 1996; Chatterji et al., 2009; Waddock, 2003)[5]. Since 1988, KLD research agency has conducted research on CSR performance, based upon publicly available information and information collected from companies. Particularly, KLD collects data from direct communication with company managers, research partners around the world, the media, public documents, and governmental and non-governmental organizations. The KLD research database is Socrates: the Corporate Social Ratings MonitorSM that contains complete sustainability reports on the largest 3000 U.S. listed companies by market capitalization. KLD is located in Boston. Since 1991, KLD has expanded its coverage several times. Initially, KLD's list of companies includes all members of the S&P 500 Index and Domini 400 Index. Members of Russell 1000 Index were included in KLD's database in 2001, and members of Russell 2000 Index were added in 2003. KLD ratings are known in social investing circles (Chatterji and Toffel, 2010; Chatterji et al., 2009). They were acquired by RiskMetrics in 2009, which was later acquired by MSCI in 2010. MSCI is a provider of stock market indexes and risk and performance analytics. Annual company environmental ratings data for KLD environmental strengths and concerns indicators were obtained from the KLD STATS database. The seven environmental strengths indicators are as follows: beneficial product and services; pollution prevention; recycling; clean energy; management systems; property, plant, and equipment; and, other strengths. The seven environmental concerns indicators are as follows: hazardous waste; regulatory problems; ozone depleting chemicals; substantial emissions; agricultural chemicals; climate change; and, other concerns. Separate KLD dimensions of environmental strengths and environmental concerns are used since the plan is to test convergent validity underlying KLD ratings. The aggregate scores of environmental strengths and environmental concerns are derived by adding the individual indicators. This adding method implies that each indicator is equally important: that of environmental strengths as well as environmental concerns. The KLD dimensions are annually

[5] http://www.kld.com/research/index.html

rated based upon binomial scale with values either 0 (bad performance) or 1 (good performance). The KLD data set consists of 5660 American companies that were rated between 1991-2008.

GES Investment Services in Scandinavia is listed among the top three research agencies in Europe (Schäfer et al., 2006)[6]. Under the name *Caring Company*, it was established in 1992 and was later renamed *Global Ethical Standard* in 2003. The headquarters are located in Sweden, Denmark, Poland, and Switzerland. The GES risk rating includes approximately 1800 MSCI World index companies, which is issued twice a year: in June and December. Ratings are based upon information that is obtained from official company documents (including annual and interim reports) and through a direct dialogue that comes in the form of surveys or site visits. Evaluation also uses public information by non-governmental organizations (NGOs), the media, and the international network of analysts in the SiRi Company Ltd. The GES database comprises a general and a specific environmental risk rating. The general environmental risk rating (environmental industry risk) reflects the environmental risk of the company's industry. The specific environmental risk (company-specific environmental index) indicates the particular environmental risk of a given company. The company-specific environmental index is based upon two sub-scores: environmental performance and environmental preparedness. Preparedness represents reputational benefits from a company's environmental policy, management systems, and regular reporting. Performance covers how a company handles environmental impacts and risks: in terms of product performance, energy use, GHG and VOC emissions, waste treatment, and other initiatives. The dimensions are assessed using a seven-point non-numerical scale from major strength (A) to major weakness (C). Companies obtain environmental ratings from Aa to Cc (capital letters indicate inherent industry risk level; lower case levels indicate company-specific risk). The GES systematic screening evaluates companies' present environmental status and readiness for the future. For empirical analysis, the seven GES Investment Services non-

[6] http://www.ges-invest.com/

numerical ratings are converted into numerical environmental scores in which companies with the lowest industry risk (A) receive a rating equal to zero and those with the highest industry risk (C) receive a rating of six. Company-specific environmental index conversions are based upon a reversed scale (i.e. the highest performance-ranked (a) companies receive a rating equal to six; the lowest performance-ranked (c) companies receive a rating of zero). The GES data set consists of 5867 U.S. companies, which were rated from December 2003 to May 2009.

ASSET4 has created a database that is said to provide transparent, objective, and auditable extra-financial information and offers a comprehensive platform for establishing benchmarks for the assessment of corporate performance (Schäfer et al., 2006)[7]. ASSET4 supports the transparency of the rating methodology that facilitates understanding the process of measuring environmental performance's scores and sub-scores for each company in the aggregate index, as well as understanding aggregation rules and data sources. The research agency was founded in 2003 and the headquarters are located in Zug, Switzerland. Investors representing more than €2.5 trillion assets under management use ASSET4 data. The ASSET4 universe includes about 2800 public world companies and covers major indices: such as NASDAQ100, S&P500, FTSE350, and MSCI World. ASSET4 collects and analyzes data from company reports, company websites, NGO websites, newspapers, journals, and trade publications. Thomson Reuters acquired ASSET4 in 2009. Annual aggregate environmental ratings data for the ASSET4 environmental pillar are used in this study. The environmental pillar measures a company's impact upon living and non-living natural systems and ecosystems, and reflects how a company uses management practices in order to generate long-term shareholder value; it is based upon three sub-dimensions: resource reduction, emission reduction, and product innovation - and upon a numerical scale ranging from 100% (good performance) to 0% (bad performance). Several narrow environmental metrics are also obtained, such as total CO2 emissions and size-adjusted CO2 emissions. Environmental

[7] http://www.asset4.com/

regulations highly influence emissions, especially in vulnerable industries. The ASSET4 data set consists of 2310 U.S. companies, which were scored between 2002-2008.

Following the aggregation of samples by company ISIN code, company name and year, the study yielded 480 U.S. MSCI World companies whose corporate environmental performance had been rated by KLD, GES, and ASSET4 at least once during the period of 2003-2008. All companies in the sample are large and publicly traded, representing a wide variety of industries (Table 1). The sample is an unbalanced panel consisting of 2880 company-year observations for the period of study. The number of observations varies depending upon which variables have missing data points.

INSERT TABLE 1 HERE

4. Results

4.1 Descriptive Statistics and Univariate Tests

Table 2 presents descriptive statistics and a correlation matrix for the sample of 480 U.S. MSCI World companies in panels A and B, respectively. The distribution of environmental measures shown in Panel A does not display considerable dispersion and markedly non-normality in the data (with the influence of some extreme values).

Panel B in Table 2 presents the Pearson Correlation Coefficients and Spearman Rho's correlations. After environmental variables have been transformed into rank-orders, non-parametric Spearman correlation coefficients are computed. Pairwise correlation coefficients are calculated among the environmental metrics in order to determine the extent to which they are related. The KLD environmental strengths dimension in this table is most highly correlated with GES environmental index (r = 0.60). Companies with high KLD environmental strengths rankings generally have higher GES environmental index ranks; those with low KLD environmental strengths rankings have lower GES environmental index rankings ($\rho = 0.62$). The GES environmental index is the most highly correlated with ASSET4 environmental pillar ($r = 0.73$). Companies that rank high in the GES environmental index generally have higher ASSET4 environmental pillar ranks when compared to other companies; those with low GES environmental index rankings have lower ASSET4 environmental pillar rankings ($\rho = 0.72$). In all of these cases, the correlations explain more than 50% of the variance in either variable. Given the fact that Pearson and Spearman correlation coefficients are very similar and different from zero, the finding suggests that the variables are consistently correlated; there is also an indication of linear relations. The results show that aggregate environmental metrics (such as KLD environmental strengths, GES environmental index, and ASSET 4 environmental pillar) seem to capture at least a substantial part of the environmental performance construct; hence, they can be considered valid. Therefore, the conclusion is that KLD environmental strengths, GES environmental index, and ASSET4 environmental pillar are consistent aggregate metrics and converge upon the same environmental performance construct.

In contrast to KLD environmental strengths, KLD environmental concerns dimension shown in Panel B correlates most highly with GES environmental industry risk dimension ($r = 0.57$) and ASSET4 one-dimensional indicator of CO_2 emissions ($r = 0.67$). Companies that rank high in KLD environmental concerns generally have higher GES environmental industry risk/ASSET4 CO_2 emissions ranks when compared to other companies; those with low KLD environmental concerns rankings have lower GES environmental industry risk/ASSET4 CO_2 emissions rankings ($\rho = 0.60$ and $\rho = 0.67$, respectively). Pearson and Spearman correlation coefficients are very similar and significant at $p < 0.001$; this suggests that these variables are consistently correlated. These results show that aggregate environmental metrics (such as GES environmental industry risk and KLD environmental concerns) are likely to capture the environmental industry construct; hence, they can be considered valid. This result is in line with previous studies. Chatterji et al. (2009) concluded that KLD environmental concerns provide transparency about pollution levels and regulatory compliance, while KLD environmental strengths do not predict pollutions levels or compliance violations. Mattingly and Berman (2006) found that KLD environmental concerns weighs upon the institutional factor much more heavily than does KLD environmental strengths. However, the authors did not differentiate between the environmental concerns and environmental strengths metrics, arguing for high correlation between variables. Prior evidence indicates that industries with high pollution propensity are more likely to provide high level of environmental performance and a high extent of environmental disclosures (Semenova and Hassel, 2009; Clarkson et al., 2008). In light of this discussion, the study concludes that KLD environmental concerns and GES environmental industry risk are consistent aggregate metrics of environmental industry construct and are, therefore, drivers of corporate environmental performance.

INSERT TABLE 2 HERE

Further analysis has been conducted to assess whether these environmental metrics are consistent across companies in different industries. Table 3 reports the results of t-tests for differences among the means of the environmental metrics for companies operating in polluting as well as green industries. The companies were divided into two sub-samples, based upon GES environmental industry risk. The first sub-sample contains of 117 companies operating in high-risk polluting industries that are scored on the basis of five to six: for example, oil & gas, industrials, and utilities. The second sub-sample includes 209 companies operating in low-risk green industries that are rated from zero to one: for example, financials, consumer services, and health care. Tests for differences in mean values statistically compare the environmental metrics of the two sub-samples; results are presented in Panel C of Table 3. The t-test indicates that the environmental ratings of companies in polluting industries are significantly higher than the environmental ratings of companies in green industries, as scored by KLD, GES, and ASSET4[8]. The non-parametric Wilcoxon rank-sum (Mann-Whitney) test, which is also shown in Table 3, Panel C, produced similar results. Based upon the dimensions of KLD environmental strengths, KLD environmental concerns, GES environmental index and ASSET4 environmental pillar, the population of companies in polluting industries once again differs from the population of companies in green industries. As aforementioned, variables are converted into ranks in order to perform the Wilcoxon rank-sum test. The Wilcoxon rank-sum tests indicate that the null hypothesis of equal population medians can be rejected because of $p < 0.001$. Hence, the conclusion is that companies in polluting and green industries are rated significantly different and consistently across all environmental metrics[9].

INSERT TABLE 3 HERE

4.2 Multivariate Tests of Hypothesis 1

The multivariate analysis in Table 4 is performed by estimating linear panel data regressions with different specifications of dependent and independent variables, as specified in Equations (1) – (3). The use of panel data requires controlling for a correlation in the error term of the regression models over time for a given company (Cameron and Trivedi, 2005; Petersen, 2009). Ignoring this panel data problem could lead to underestimated standard errors and inflated *t*-statistics. The parameters of the model are computed by using techniques suggested in (Petersen, 2009) with cluster-robust standard errors that correct serial correlation and heteroskedasticity. The advantage of the panel data approach is the possibility of a consistent estimation of the model, which controls bias from omitted variables. For the panels used in this study, the econometric models

[8] Without assuming equal variances, untabulated results for t-test are similar to those that rest upon an equal-variances assumption.
[9] Untabulated results for t-test performed by dividing the companies based upon the ICBIN sectors are consistent with the results from GES Industry risk estimates.

are as follows: pooled, fixed effects, and random effects models. The fixed effects model allows the unobserved factors to be correlated with regressors and permits the identification of the marginal effect for time-varying variables. The random effects model treats any unobserved individual heterogeneity as being distributed independently of the repressors (Cameron and Trivedi, 2005). Since this study estimates the individual-specific effects panel data models, the results of the Lagrangian multiplier test and the Hausman test show that the fixed effects model is the appropriate choice for the data (Greene, 2003). To ensure that the regression results are not unduly sensitive to outliers, this study excludes observations identified using *DFITS* statistics proposed by Belsley et al. (1980)[10]. For the purpose of the regressions reported below, the independent regressors, whilst showing indications of high collinearity in Table 2, have no pairwise correlation coefficients in excess of 80 percent. This indicates that the threat of multicollinearity is limited[11]. In addition, calculations of VIF statistics are reported in all specifications[12]. Coefficients on dependent environmental variables are reported based upon their one-tailed tests of significance.

The results summarized in Table 4 indicate that the coefficients for KLD environmental strengths, ASSET4 environmental pillar, and GES environmental index are significantly positive in all tests. The measure of KLD environmental concerns is significantly positively associated with GES environmental index and KLD environmental strengths. Apart from the univariate tests, the results of the multivariate analysis do not allow the distinguishing between two constructs (such as environmental performance and environmental industry risk) and are potentially sensitive to the difference in the measurement scales of environmental metrics. Overall, the multivariate analysis suggests that once unobservable company heterogeneity is

---

[10] As Belsley, Kuh and Welsch (1980) recommend, a cut-off value of $|DFITS_j| > 2\sqrt{k/N}$ indicates highly influential observations, where k is the number of estimated coefficients, *n* is sample size, *j* is an observation.

[11] As rule of thumb, correlation between explanatory variables in excess of 0.8 suggests that multicollinearity is a serious problem (Gujarati, 1995).

[12] As common cut-off criterion, a mean VIF of the factors considerably larger than 1 suggests multicollinearity (Chatterjee and Price, 1991).

controlled for, the environmental performance metrics are significantly positively related. The study finds empirical support for hypothesis H1 for KLD environmental strengths, GES environmental index, and ASSET4 environmental pillar.

INSERT TABLE 4 HERE

4.3 Multivariate Tests of Hypothesis 2

Table 5 provides the results of regression models based upon Equation (4). Columns of the panels report coefficients on dependent variable GES Environmental index and their one-tailed test of significance. The study uses two proxies for industry risk to test Equation (4): GES environmental industry risk and KLD environmental concerns, and two alternative proxies for profitability: ROA and ROE. Panel A and Panel B show that GES environmental industry risk and KLD environmental concerns are significantly positively associated with corporate environmental performance (β_1 = 0.09, *t*-value = 3.59 and β_1 = 0.71, *t*-value = 18.06, respectively). These results are not sensitive to the choice of proxy for profitability and are consistent with Chatterji and Toffel (2010): that KLD environmental concerns lead to improvements in environmental performance. Company size has a significantly positive relation to environmental performance for both GES and KLD industry risk metrics models and profitability proxies. This result is consistent with previous findings by Artiach et al. (2010). ROA is significantly positively associated with environmental performance for both GES and KLD models. ROE is significantly positively associated with environmental performance in the GES model (β_3 = 0.14, *t*-value = 1.94); however, this effect is insignificant in KLD model. Therefore, the multivariate analysis confirms a positive relationship between environmental performance and industry risk and company size. This study finds empirical support for hypothesis H2. Interestingly, the relationship between environmental performance and profitability is sensitive to both the specification of the model and the choice of proxy for profitability. On this basis, the conclusion is that industry risk drives companies to active environmental work[13].

INSERT TABLE 5 HERE

4.4 Additional Analysis

The foundation for additional empirical test is the regression of GES environmental index on GES environmental industry risk, emissions, and profitability. The initial sample has been reduced due to missing values of the emissions variable. The emissions variable scaled by total sales revenue is typically referred to as a measure of pollution performance (Jaggi and Freedman, 1992) or as a measure of environmental efficiency or *eco-efficiency* (Chatterji and Toffel, 2010). The study uses two alternative proxies for profitability: ROA and ROE; the results are shown in Panel C1 and Panel C2, respectively. GES environmental industry risk has a significantly positive relationship to GES environmental index (β_1 = 0.10, *t*-value = 2.15). Emissions scaled by sales revenue have a significantly negative relationship to GES environmental index (β_2 = -58.83, *t*-value = -3.11). These results are not sensitive to the choice of proxy for profitability. The results for the GES environmental industry risk are consistent with the aforementioned reported findings. The results for emissions are consistent with previous findings by Clarkson et al. (2008). ROA is insignificantly associated with environmental performance, while ROE is significantly negative (β_3 = -1.27, *t*-value = -2.17). The additional multivariate analysis confirms that environmental performance has a positive relationship to industry risk and a negative association with the emissions variable. Interestingly, the relationship between environmental performance and profitability is sensitive to both the specification of the model and the choice of proxy for profitability.

[13] Untabulated results for ASSET 4 environmental pillar as environmental performance proxy are similar to those presented in Table 5. Untabulated results for KLD environmental strengths as environmental proxy indicate that GES environmental industry risk has an insignificant relationship to KLD environmental strengths for both profitability proxies. KLD environmental concerns are significantly positively associated with corporate environmental performance, proxied by KLD environmental strengths.

5. Conclusions

This study has investigated the consistency of environmental metrics compiled by KLD, GES, and ASSET4 rating services; it has also explored the association between industry-specific factors and corporate environmental performance for U.S. companies from MSCI World universe 2003-2008. The study has three main results. First, multi-dimensional environmental performance ratings and rankings are highly related and consistent across companies and industries; specifically, Pearson and Spearman correlation coefficients explain more that 50% of variance in KLD environmental strengths, GES environmental index, and ASSET4 environmental pillar metrics. The univariate t-test revealed that the environmental performance of companies from polluting industries gets significantly higher scores across KLD, GES, and ASSET4 metrics than the environmental performance of companies from green industries. The multivariate model explained a significant amount of variance in the ratings. Secondly, the KLD environmental concerns metric is likely to be a good proxy for environmental risk, summarizing the effect of industrial activities. The KLD concerns are highly correlated to GES environmental industry risk and ASSET4 CO_2 emissions with “r” and “ρ” values over 50%. Third, the environmental industry risk, impact or concerns must be considered as factors that drive corporate environmental performance. Industry-specific risk and company-specific concerns factors are positively related to corporate environmental performance in GES, ASSET4, and KLD models. The results discussed have implications on a large body of literature that finds evidence on the link between market value and environmental performance by utilizing multi-dimensional environmental metrics and a scope of studies that argues it is industry risk that drives companies to active environmental work.

The main contribution of this paper is a validation of corporate environmental metrics. The research focuses upon two distinct constructs: the impacts of industry-specific activities (industry risk) and the impacts of company-specific activities (environmental performance); this bears resemblance to the environmental concept discussed in prior literature. The research indicates that KLD environmental strengths and KLD environmental concerns fail to converge on the same environmental performance construct. A key observation is that combining non-convergent environmental metrics in the construction of a composite performance measure or in modelling settings can mask an underlying relationship between variables and, therefore, confuse the interpretation of observed relations. Additionally, one contribution of this study is the focus upon environmental performance drivers on the company level: particularly, the identification of environmental industry-related factors. Examples of drivers of corporate environmental performance are GES environmental industry risk and KLD environmental concerns.

The implications of results of the paper are summarized in this paragraph. First, the three environmental performance ratings do a reasonable consistent job of aggregating an underlying environmental performance construct. High correlations among environmental performance variables and large amount of explained variance show that evaluation methods generalize different types of environmental indicators and environmental information with low potential noise; this is consistent with Hypothesis 1. Consistent and reliable environmental metrics assist users in making informed decisions upon the basis of environmental performance, enabling the differentiation between environmental drivers and environmental outcomes in existing rating systems. Secondly, leading corporate environmental performance companies are most likely to operate in industries with high environmental risk. Companies in high risk or polluting industries are more communicative and transparent because they are pressured and draw the attention of a wide range of NGO environmental groups, state regulators, journalists, and an increasing number of investors and asset managers as stakeholders; this is consistent with the arguments underlying Hypothesis 2. Passive environmental management and disclosure are less feasible for companies in polluting industries because, when compared with companies in green industries, they correspondingly create larger environmental risks and are subject to more scrutiny from both environmental regulators and stakeholders. Environmental performance leaders are most likely to be large companies in polluting industries that have better availability to resources and lower average costs in achieving high environmental performance; this is consistent with prior literature on ratings.

This study may be extended in several ways. Further research is needed to develop a classification of disaggregated environmental performance indicators and demonstrate their construct validity. The analysis of environmental indicators may also shed light upon their operationalisation in aggregate environmental performance measures. While this study is limited by KLD, GES, and ASSET4 environmental metrics, additional composite environmental variables may provide more confidence in the validity of environmental performance metrics. Overall, this study is an important contribution to the ongoing debate about the consistency of environmental metrics, the determinants of corporate environmental performance, and the usefulness of environmental corporate ratings.

References

Azzone, G., Noci, G., 1996. Defining environmental performance indicators: an integrated framework. Business Strategy and the Environment 5, 69-80.

Artiach, T., Lee, D., Nelson, D., Walker, J., 2010. The determinants of corporate sustainability performance. Accounting and Finance 50 (1), 31-51.

Bagozzi, R., Yi, Y., Phillips, L., 1991. Assessing construct validity in organizational research. Administrative Science Quarterly 36, 421-458.

Baron, D., Harjoto, M., Jo, H., 2009. The economics and politics of corporate social performance. Research Paper Series, Stanford, the Graduate School of Business.

Barth, M., McNichlos, M., 1994. Estimation and market valuation of environmental liabilities relating to superfund sites. Journal of Accounting Research 32 (Supplement), 177-209.

Bassen, A., Kovacs, A., 2008. Environmental, social and governance key performance indicators from a capital market perspective. The Journal for Business, Economics & Ethics 9/2, 182-192.

Besley, D., Kuh, E., Welsch, R., 1980. Regression Diagnostic. New York: Wiley.

Berman, L., Wicks, A., Kotha, S., Jones. T., 1999. Does stakeholder orientation matter? The relationship between stakeholder management models and firm financial performance. The Academy of Management Journal 42 (5), 488-506.

Blacconiere, W., Northcut, W., 1997. Environmental information and market reactions to legislation. Journal of Accounting, Auditing and Finance 12 (2), 149-78.

Brown, B., Perry, S., 1994. Removing the financial performance halo from Fortune's "most admired" companies. Academy of Management Journal 37, 1347-1359.

Cameron, C., Trivedi, P., 2005. Microeconometrics, methods and applications, Cambridge University Press, New York.

Chatterjee, S., Price, B., 1991. Regression Analysis by Example, 2nd edn (New Yourk: Wiley).

Chatterji, A., Levine, D., 2006. Breaking down the wall of codes: evaluating non-financial performance measurement. California Management Review 48 (2), 29-51.

Chatterji, A., Levine, D., Toffel, M., 2009. How well do socially ratings actually measure corporate social responsibility? Journal of Economics and Management Strategy 18 (1), 125-169.

Chatterji, A., Toffel, M., 2010. How firms respond to being rated. Strategic Management Journal, Published online Early View in Wiley InterScience (www.interscience.wiley.com).

Cho, C., Patten, D., 2007. The role of environmental disclosures as tools of legitimacy: a research note. Accounting, Organizations and Society 32, 639-647.

Clarksson, P., Li, Y., Richardsson, G., Vasvari, F., 2006. Does it really pay to be green? Determinants and consequences of pro-active environmental strategies. Working Paper (http://kcjd.shufe.edu.cn/News/UploadFiles/news/news-2007-5-23-001.pdf).

Clarkson, P., Li, Y., Richardson, G., Vasvari, F., 2008. Revising the relation between environmental performance and environmental disclosure: en empirical analysis. Accounting, Organizations and Society 33, 303-327.

Derwall, J., Guenster, N., Bauer, R., Koedijk, K., 2005. The eco-efficiency premium puzzle. Financial Analysts Journal 61(2), 51-63.

Elsayed, K., Paton, D., 2005. The impact of environmental performance on firm performance: static and dynamic panel data evidence. Structural Change and Economic Dynamics 16, 395-412.

Entine, J., 2003. The myth of social investing. Organization & Environment 16 (3), 352-368.

Fryzell, G., Wang, J., 1994. The Fortune corporate 'reputation' index: reputation for what? Journal of Management 20 (1), 1-14.

Jaggi, B., Freedman, M., 1992. An examination of the impact of pollution performance on economic and market performance: pulp and paper firms. Journal of Business Finance and Accounting 19 (5), 697-713.

Johnson, R., Greening, D., 1999. The effects of corporate governance and institutional ownership types on corporate social performance. The Academy of Management Journal 42 (5), 564-576.

Gamble, G., Hsu, K., Kite, D., Radtke, R., 1995. Environmental disclosures in annual reports and 10Ks: en examination. Accounting Horizons 9 (3), 34-64.

Gerde, V., Logsdon, J., 2001. Measuring environmental performance: use of the Toxic Release Inventory (TRI) and other US environmental databases. Business Strategy and the Environment 10, 269-285.

Greene, W. H. (2003) *Econometric Analysis*. (New Jersey: Prentice-Hall).

Griffin, J., Mahon, J., 1997. The corporate social performance and corporate financial performance debate. Business & Society 36 (1), 5-31.

Gujarati, D., 1995. Basic Econometrics, 3rd edn (McGraw-Hill, Singapore).

Guenster, N., Derwall, J., Bauer, R. and Koedijk, K. (2006) The economic value of corporate eco-efficiency, Working paper (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=675628), RSM Erasmus University, Maastricht University, Tilburg University and European Centre for Corporate Engagement.

Hamilton J., 1995. Pollution as news: media and stock market reactions to the Toxics Release Inventory data. Journal of Environmental Economics and Management 28, 98–113.

Hart, S., Ahuja, G., 1996. Does it pay to be green? An empirical examination of the relationship between emission reduction and firm performance. Business Strategy and the Environment 5, 30-37.

Hillman, A., Keim, G., 2001 Shareholder value, stakeholder management, and social issues: what's the bottom line? Strategic Management Journal 22 (2), 125-139.

Ilinitch, A., Schaltegger, S., 1995. Developing a green business portfolio. Long Range Planning 28(2), 29-38.

Ilinitch, A., Soderstrom, N., Thomas, T., 1998. Measuring corporate environmental performance. Journal of Accounting and Public Policy 17, 383-408.

Khanna, M., Quimio, W., Bojilova, D., 1998. Toxic release information: a policy tool for environmental protection. Journal of Environmental Economics and Management 36, 243-266.

Kempf, A., Osthoff, P., 2007. The effect of socially responsible investing on portfolio performance. European

Financial Management 13 (5), 908-922.
King, A., Lenox, M., 2001. Does it really pay to be green? An empirical study of firm environmental and financial performance. Journal of Industrial Ecology 5 (1), 105-116.
King, A., Lenox, M., 2002. Exploring the locus of profitable pollution reduction. Management Science 48 (2), 289-299.
Konar, S., Cohen, M., 2001. Does the market value environmental performance? The Review of Economics and Statistics 83 (2), 281-289.
Lober, D., 1996. Evaluating the environmental performance corporations. The Journal of Managerial Issues 8(2), 184-205.
Margolis, J., Elfenbein H., Walsh, J., 2007. Does it pay to be good? A meta-analysis and direction of research on the relationship between corporate social and financial performance. Working Paper. Harvard University. (http://stakeholder.bu.edu/2007/Docs/Walsh,%20Jim%20Does%20It%20Pay%20to%20Be%20Good.pdf)
Mattingly, J., Berman, S., 2006. Measurement of corporate social action. Discovering taxonomy in the Kinder Lydenburg Domini ratings data. Business & Society 45 (1), 20-46.
Olsson, R., 2007. Portfolio performance and environmental risk. Sustainable Investment Research Platform Working Papers (www.sirp.se).
Patten, D., 2002. The relation between environmental performance and environmental disclosure: a research note. Accounting, Organizations and Society 27, 763-773.
Petersen, M. A., 2009. Estimating standard errors in financial panel data sets: comparing approaches. Review of Financial Studies 22, 435-480.
Scholtens, B., Zhou, Y., 2008. Stakeholder relations and financial performance. Sustainable Development 16, 213-232.
Schäfer, H., 2005. International corporate social responsibility systems. Conceptual outline and empirical results. Journal of Corporate Citizenship 20 (Winter), 107-120.
Schäfer, H., Beer, J., Zenker, J., Fernandes, P., 2006. Who is who in corporate social responsibility rating. A survey of internationally established rating systems that measure corporate responsibility. Bertelsmann Foundation and University Stuttgart (http://www.global-ethic-now.de/gen-eng/0d_weltethos-und-wirtschaft/0d-pdf/04-verantwortung/transparenzstudie_bertelsmann.pdf)
Sharfman, M., 1996. The construct validity of the Kinder, Lydenberg & Domini social performance ratings data. Journa of Business Ethics 15, 287-296.
Semenova, N., Hassel, L., 2008. Financial outcomes of environmental risk and opportunity for US companies. Sustainable Development 16, 195-212.
Sinkin, C., Wright, C., Burnett, R., 2008. Eco-efficiency and firm value. Journal of Accounting and Public Policy 27, 167-176.
Stanwick, P., Stanwick, S., 1998. The relationship between corporate social performance and organizational size, financial performance, and environmental performance: an empirical examination. Journal of Business Ethics 17, 195-204.
Statman, M., Glushkov, D., 2009. The wages of social responsibility. Financial Analysts Journal 65 (4), 33-46.
Tyteca, D., 1996. On the measurement of the environmental performance of firms – a literature review and a productive efficiency perspective. Journal of Environmental Management 46, 281-308.
Turban, D., Greening, D., 1997. Corporate social performance and organizational attractiveness to prospective employees. The Academy of Management Journal 40 (3), 658-672.
Ziegler, A., Schröder, M., Rennings, K., 2007. The effect of environmental and social performance on the stock performance of european corporations. Environmental and Resource Economics 37, 661-680.
Waddock, S., 2003. Myths and realities of social investing. Organizations and Environment 16 (3), 369-380.
Waddock, S., Graves, S., 1997. The corporate social performance-financial performance link. Strategic Management Journal 18 (4), 303-319.
William, Y., Welford, R., 1998. An environmental performance measurement framework for business. Greener Management International 21 (Spring), 30-50.
Wood, D., 1991. Corporate social performance revisited. Academy of Management Review 16 (4), 691-718.

Table 1. Distribution of Companies Across Industries and Stock Exchanges

Panel A: Industry Classified Companies Based Upon the Industry Classification Benchmark

ICBIN Code	Sector	Number of Companies
8000	Financials	96
5000	Consumer Services	74
2000	Industrials	63
9000	Technology	58
3000	Consumer Goods	56
4000	Health Care	48
7000	Utilities	30
0001	Oil & Gas	27
1000	Basic Materials	20
6000	Telecommunications	8
	Total	480

Panel B: Industry-Risk Classified Companies Based Upon the GES Investment Service Rating

GES Industry Risk	Number of Companies
High (rated 5,6)	117
Medium (rated 2,3,4)	326
Low (rated 0,1)	209
Total	480

Panel C: Companies Classified According to Stock Exchange

Stock Exchange	Number of Companies
NYS	392
NAS	85
ASE	3
Total	480

Table 2. Descriptive Statistics and Correlation Matrix

Panel A: Descriptive Statistics

Variable	Mean	SD	Min	Max	Skew	Kurt
1. KLD Env-Str	0.37	0.79	0	4	2.43	5.82
2. KLD Env-Con	0.57	1.08	0	5	2.00	3.34
3. GES Env-Industry	2.89	2.17	0	6	0.13	-1.41
4. GES Env-Index	1.32	1.61	0	6	0.93	-0.12
5. A4 Env-Pillar	0.42	0.28	0.10	0.98	0.74	-0.97
6. Log A4 Env-CO2	14.55	2.15	5.52	18.92	-0.17	0.12

Note: 2880 company-year observations.

Panel B: Correlation Matrix

	KLD Env-Str	KLD Env-Con	GES Env-Industry	GES Env-Index	A4 Env-Pillar	Log A4 Env-CO2
1. KLD Env-Str	1.00	*0.42 (0.00)*	*0.39 (0.00)*	*0.62 (0.00)*	*0.51 (0.00)*	*0.13 (0.00)*
	n=2739	*n=2739*	*n=2466*	*n=2404*	*n=2461*	*n=519*
2. KLD Env-Con	0.41 (0.00)	1.00	*0.60 (0.00)*	*0.50 (0.00)*	*0.44 (0.00)*	*0.67 (0.00)*
	n=2739	n=2739	*n=2466*	*n=2404*	*n=2461*	*n=519*
3. GES Env-Industry	0.33 (0.00)	0.57 (0.00)	1.00	*0.51 (0.00)*	*0.40 (0.00)*	*0.66 (0.00)*
	n=2466	n=2466	n=2515	*n=2435*	*n=2270*	*n=519*
4. GES Env-Index	0.60 (0.00)	0.46 (0.00)	0.47 (0.00)	1.00	*0.72 (0.00)*	*0.36 (0.00)*
	n=2404	n=2404	n=2435	n=2436	*n=2222*	*n=508*
5. A4 Env-Pillar	0.54 (0.00)	0.45 (0.00)	0.37 (0.00)	0.73 (0.00)	1.00	*0.21 (0.00)*
	n=2461	n=2461	n=2270	n=2222	n=2502	*n=524*
6. Log A4 Env-CO2	0.09 (0.05)	0.67 (0.00)	0.64 (0.00)	0.28 (0.00)	0.22 (0.00)	1.00
	n=519	n=519	n=517	n=508	n=524	n=524

Note: 2880 company-year observations.

Panel A presents the descriptive statistics for the pooled sample of 480 U.S. MSCI World companies, which were rated between 2003-2008.
Panel B presents the Pearson correlation coefficients and Spearman rho pairwise correlations.
(Spearman rho values are reported above the diagonal and Pearson correlation coefficients are reported below. Correlation coefficients are reported with the p-value using a two-tailed test of significance shown in parentheses).

Table 3. Univariate Tests

Variable	Mean	SD	Min	Max	Skew	Kurt	Num. Obs.
Panel A: High GES Env-Industry Risk (rated ≥5), 702 Company-Year Observations							
1. KLD Env-Str	0.79	0.99	0	4	1.27	1.07	681
2. KLD Env-Con	1.72	1.42	0	5	0.45	-0.73	681
4. GES Env-Index	2.28	1.50	0	6	0.02	-0.40	624
5. A4 Env-Pillar	0.56	0.27	0.10	0.97	-0.01	-1.45	614
Panel B: Low GES Env-Industry Risk (rated ≤1), 1254 Company-Year Observations							
1. KLD Env-Str	0.09	0.38	0	3	4.98	27.09	1188
2. KLD Env-Con	0.09	0.40	0	4	5.58	34.92	1188
4. GES Env-Index	0.49	1.01	0	6	2.21	4.65	1049
5. A4 Env-Pillar	0.31	0.21	0.10	0.97	1.62	1.47	1085

Panel C: Difference in Mean/Median Values for Companies in Polluting Industries and Companies in Green Industries

	t-test			Wilcoxon Rank-Sum Test	
	Mean diff	t-stat	p-value	z-stat	p-value
1. KLD Env-Str	0.70	21.74	0.00	21.78	0.00
2. KLD Env-Con	1.63	37.30	0.00	31.22	0.00
4. GES Env-Index	1.79	29.17	0.00	24.37	0.00
5. A4 Env-Pillar	0.25	21.12	0.00	19.18	0.00

Panel A presents the descriptive statistics for the environmental variables using the pooled sample of 117 companies operating in polluting industries.
Panel B presents the descriptive statistics for the pooled sample of 209 companies operating in green industries. The companies in Panel A and Panel B are divided into the polluting industry and green industry sub-sets based on GES Env-Industry risk scores.
Panel C presents the t-test and Wilcoxon rank-sum (Mann-Whitney) test for differences in mean/median variables for companies operating in polluting industries and companies operating in green industries.

Table 4. Multivariate Tests

Variable	Linear Panel Data Regressions **Panel A: Dependent Variable: GES Env-Index**				Linear Panel Data Regressions **Panel B: Dependent Variable: A4 Env-Pillar**			
	VIF	Pooled OLS	*Fixed Effects (Within) OLS*	Random Effects GLS	VIF	Pooled OLS	*Fixed Effects (Within) OLS*	Random Effects GLS
Intercept		-0.43***	*0.44****	-0.17***		0.25***	*0.32****	0.26***
		(-12.33)	*(6.38)*	(-4.04)		(45.01)	*(28.53)*	(47.55)
KLD Env-Str	1.44	0.52***	*0.20****	0.46***	1.72	0.04***	*0.04****	0.05***
		(14.57)	*(4.15)*	(12.62)		(5.87)	*(4.86)*	(8.85)
KLD Env-Con	1.28	0.20***	*0.14****	0.24***	1.36	0.02***	*0.01*	0.02***
		(7.66)	*(2.28)*	(9.44)		(3.67)	*(0.38)*	(4.45)
A4 Env-Pillar	1.56	3.12***	*1.32****	2.46***				
		(29.26)	*(8.77)*	(22.77)				
GES Env-Index					1.83	0.13***	*0.07****	0.12***
						(33.24)	*(10.20)*	(30.69)
Adj. R^2		0.71	*0.71*	0.73		0.68	*0.67*	0.70
LM (BP)		48.29***	*48.29****	48.29***		25.05***	*25.05****	25.05***
Hausman		434.11***	*434.11****	434.11***		115.73***	*115.73****	115.73***
Mean VIF	1.43				1.64			
Num. Obs.	2008	2008	*2008*	1975	2070	2070	*2070*	2025

The table shows the outcome of estimating linear panel regressions across different specifications of GES, ASSET4, and KLD environmental performance metrics. The unbalanced panel contains 480 companies constituting 2880 company-year observations.
The table reports pooled OLS, fixed effects (within) OLS and random effects GLS coefficients with the cluster-robust t-statistic (in parentheses).
Significance at the 1%, 5% and 10% level is indicated by ***, **, and *, respectively (one-tailed tests).
LM (BP) is the Breusch and Pagan Lagrangian multiplier test for random effects. Hausman is the Hausman test for fixed effects over random effects. VIF is variance inflator factor test for multicollinearity.
Approximately 6% of the observations in Panel A are flagged by the DFITS cut-off criterion (Belsley, Kuh, and Welsch, 1980; Baum, 2006) as highly influential.
Approximately 4% of the observations in panel B are flagged by the DFITS cut-off criterion (Belsley, Kuh, and Welsch, 1980; Baum, 2006) as highly influential.

Table 4 (continued). Multivariate Tests

Variable	Linear Panel Data Regressions **Panel C: Dependent Variable: KLD Env-Str**			
	VIF	Pooled OLS	*Fixed Effects (Within) OLS*	Random Effects GLS
Intercept		-0.11***	*0.004*	-0.07***
		(-4.26)	*(0.11)*	(-3.84)
KLD Env-Con	1.24	0.25***	*0.20****	0.16***
		(6.53)	*(3.83)*	(6.97)
A4 Env-Pillar	2.09	0.31***	*0.27****	0.32***
		(3.61)	*(3.86)*	(5.60)
GES Env-Index	2.23	0.17***	*0.06****	0.13***
		(8.82)	*(4.24)*	(10.85)
Adj. R^2		0.44	*0.42*	0.47
LM (BP)		902.93***	*902.93****	902.93***
Hausman		66.79***	*66.79****	66.79***
Mean VIF	1.85			
Num. Obs.	2008	1960	*1975*	1942

Approximately 7% of the observations in panel C are flagged by the DFITS cut-off criterion (Belsley, Kuh, and Welsch, 1980; Baum, 2006) as highly influential.

Table 5. Multivariate Tests

Variable	Linear Panel Data Regressions **Dependent variable: GES Env-Index**					
	Panel A1: Fixed Effects Model	Panel A2: Fixed Effects Model	Panel B1: Pooled OLS	Panel B2: Pooled OLS	Panel C1: Pooled OLS	Panel C2: Pooled OLS
Intercept	-6.80***	-6.64***	-1.84***	-0.89**	2.91***	2.88***
	(-6.45)	(-6.52)	(-3.16)	(-1.72)	(11.66)	(11.11)
GES Env-Industry	0.09***	0.08***			0.10**	0.11***
	(3.59)	(3.26)			(2.15)	(2.21)
KLD Env-Con			0.71***	0.70***		
			(18.06)	(17.84)		
A4 Env-CO2/Sales revenue					-58.83***	-61.37***
					(-3.11)	(-3.30)
LnTA	0.47***	0.47***	0.15***	0.10***		
	(7.47)	(7.48)	(4.37)	(3.22)		
ROA	2.21***		5.09***		-0.50	
	(2.73)		(3.18)		(-0.24)	
ROE		0.14**		0.11		-1.27**
		(1.94)		(0.53)		(-2.17)
Adj. R^2	0.14	0.11	0.30	0.29	0.02	0.03
LM (BP)	129.02***	119.44***	0.22	0.38	0.06	0.65
Hausman	88.76***	90.72***				
Num. Obs.	2340	2349	2273	2290	407	415

Panel A and Panel B present the outcome of estimating linear panel regressions of GES Env-Index on GES environmental industry risk, KLD environmental concerns, company size, and profitability. The unbalanced panel contains 480 companies constituting 2880 company-year observations between 2003-2008.
Panel C shows the outcome of estimating linear panel regressions of GES Env-Index on GES environmental industry risk, emissions, and profitability. The unbalanced panel contains 103 companies constituting 618 company-year observations between 2003-2008.
The table reports fixed effects (within) OLS and pooled OLS coefficients with the cluster-robust t-statistic (in parentheses).
Significance at the 1%, 5% and 10% level is indicated by ***, **, and *, respectively (one-tailed tests).
LM (BP) is the Breusch and Pagan Lagrangian multiplier test for random effects. Hausman is the Hausman test for fixed effects over random effects.