07-051

How Well Do Social Ratings Actually Measure Corporate Social Responsibility?

Aaron K. Chatterji
David I. Levine
Michael W. Toffel

How Well Do Social Ratings Actually Measure Corporate Social Responsibility?

Aaron K. Chatterji
Fuqua School of Business
Duke University
1 Towerview Drive
Durham, NC 27708
ronnie@duke.edu

David I. Levine
Haas School of Business
University of California at Berkeley
545 Student Services Building #1900
Berkeley, CA 94720-1900
levine@haas.berkeley.edu

Michael W. Toffel[*]
Harvard Business School
Soldiers Field
Boston, MA 02163
mtoffel@hbs.edu

February 20, 2008

Forthcoming in *Journal of Economics & Management Strategy*

We gratefully acknowledge KLD's Peter Kinder, Institutional Shareholder Services' Joe Henzlik, and the Investor Responsibility Research Center's Jennifer Hodge and Michael Langlais for providing, and helping us to understand, the datasets. We appreciate comments from Michael Barnett, Graham Bullock, Helen Liang, Will Mitchell, Robert Salomon, Bill Simpson, Ralf Steinhauser, and participants in the 2006 Institutions and Innovation Conference at the Harvard Business School.

[*] Corresponding author

How Well Do Social Ratings Actually Measure Corporate Social Responsibility?

Abstract

Ratings of corporations' environmental activities and capabilities influence billions of dollars of "socially responsible" investments as well as some consumers, activists, and potential employees. In one of the first studies to assess these ratings, we examine how well the most widely used ratings—those of Kinder, Lydenberg, Domini Research & Analytics (KLD)—provide transparency about past and likely future environmental performance. We find KLD "concern" ratings to be fairly good summaries of past environmental performance. In addition, firms with more KLD concerns have slightly, but statistically significantly, more pollution and regulatory compliance violations in later years. KLD environmental strengths, in contrast, do not accurately predict pollution levels or compliance violations. Moreover, we find evidence that KLD's ratings are not optimally using publicly available data. We discuss the implications of our findings for advocates and opponents of corporate social responsibility as well as for studies that relate social responsibility ratings to financial performance.

How Well Do Social Ratings Actually Measure Corporate Social Responsibility?

1. Introduction

An important element of management strategy is managing how stakeholders view a company's impact on the natural environment. Stakeholders' perceptions can be critical to firm performance and sometimes even survival (Hart, 1995; Russo and Fouts, 1997; Berman et al., 1999). Many companies attempt to enhance their environmental image by mitigating deleterious effects on the environment and publicizing (with varying degrees of accuracy) their successes. It can be difficult, however, for stakeholders to evaluate companies' environmental impacts. Stakeholders are often unaware of the full range of firms' activities and lack access to or the expertise needed to analyze relevant environmental data (Lyon and Maxwell, 2006).

Social and environmental rating agencies seek to make corporations' environmental effects more transparent. These rating agencies can examine firms' *past* environmental performance and environmental management activities. In addition, they can also consider a firm's *future outlook*, such as by analyzing their environmental management plans and investments that purport to enhance future environmental performance.[1] Just as credit ratings "enhance transparency and efficiency in debt capital markets by reducing the information asymmetry between borrowers and lenders,"[2] social ratings aim to provide social investors accurate information that makes transparent the extent to which firms' behaviors are socially responsible.

Poor social and environmental ratings can harm a company's performance and reputation. For example, Kinder, Lydenberg, Domini Research & Analytics (KLD) dropped Coca-Cola Co., from its Broad Market Social Index in July 2006 because of concerns about the company's labor

and environmental practices in the developing world. As a result, TIAA-CREF, the largest U.S. retirement fund, subsequently sold more than 50 million shares of Coca-Cola Co. stock.[3] Critics of the company also seized on KLD's action as support for their longstanding complaints against Coca-Cola Co.[4]

Despite their increasing popularity, social ratings are rarely evaluated and have been criticized for their own lack of transparency. Prominent environmentalist Paul Hawken, for example, recently harshly criticized socially responsible investing, noting that "the screening methodologies and exceptions employed by most SRI [socially responsible investment] mutual funds allow practically any publicly-held corporation to be considered as an SRI portfolio company."[5]

It remains unclear whether social ratings are actually providing transparency that helps stakeholders identify environmentally responsible companies. KLD's social and environmental ratings are among the oldest and most influential and, by far, the most widely analyzed by academics. In this paper, we examine the extent to which KLD's ratings make transparent to stakeholders which companies are environmentally responsible.

Investors who seek transparency are looking for some combination of (1) accuracy in summarizing past performance, and (2) careful evaluation of current managerial actions likely to influence future environmental performance. We thus first empirically investigate the extent to which KLD's ratings accurately capture past performance. We then evaluate how well KLD's ratings measure current managerial decisions and organizational capabilities that affect future environmental performance.[6] The results presented here constitute the first rigorous evaluation of the transparency of KLD's environmental ratings.

If the prevailing social ratings are not providing reasonable transparency, the investors and other stakeholders who rely on them to identify desirable target companies might be misallocating resources. In contrast, if the ratings do make environmental performance transparent, they can assist stakeholders interested in rewarding or punishing firms on the basis of environmental performance. Furthermore, our study will inform the strategies of stakeholders who use environmental ratings to make decisions about procurement and political action (e.g., lobbying and boycotts).

Our results will also inform the substantial empirical literature that links corporate financial performance to corporate social performance. More than 100 studies have examined whether CSR metrics predict financial performance, with a variety of results (for reviews see Margolis and Walsh 2003 and Margolis, Elfenbein and Walsh 2007). Those studies finding little correlation between CSR metrics and financial performance may understate the relationship between actual CSR and financial performance if the CSR metrics are noisy indicators of true CSR activities. At the same time, those studies finding a positive correlation between CSR metrics and financial performance may overstate the relationship between actual CSR and financial performance if customers or other stakeholders are misled by the erroneous CSR metrics (for example, by successful "greenwashing" campaigns).

Others have stressed the importance of measuring how well CSR metrics provide transparency about CSR behaviors and performance. For example, scholars such as Margolis and Walsh (2003, p. 297) have argued that "we need to understand the conditions under which a corporation's efforts benefit society" before we can understand the "link between a firm's social and financial performance." More critically, Jon Entine recently posited that "evidence [on socially responsible investing] in many areas, from corporate governance, to supply chain analysis, to

energy, suggests that SRI funds are very sloppy and often flat out wrong in identifying 'doing good.'"[7]

In this paper, we empirically investigate the extent to which CSR metrics from one of the world's most prominent social rating agencies provide transparency about prior and future corporate social performance. As far as we know, our analysis is the first to examine this important topic.

2. Transparency, Investors, and the Role of Environmental Ratings

In 2005, investors in the United States allocated nearly $1.7 trillion to socially screened portfolios including mutual funds.[8] In this section we outline the diverse motives that social investors and other stakeholders may have to seek transparency about firms' past and future performance. We then explain how a social rating agency should behave to meet the overlapping requirements of investors with these several motives.

We characterize the motivations of social investors as: financial, deontological, consequentialist, and expressive (see Table 1).[9] An investor or other stakeholder can be driven by any combination of these motives. We describe how each motive leads investors to value social ratings that (1) provide transparency about historical social performance, and (2) provide transparency about current management practices that influence future social performance. For example, such management practices include information about a firm's significant investments in energy-saving technology or recent hiring of a new CEO with a reputation for environmental stewardship.

The first motive we consider is the belief that socially responsible companies will perform better financially. Prior research has examined how corporate social responsibility (CSR) can benefit companies by attracting socially responsible consumers (Bagnoli and Watts, 2003), reducing

the threat of regulation (Maxwell et al., 2000), improving their reputations with consumers (Lev et al., 2006), and reducing concern from activists and non-governmental organizations (Baron, 2001; Lyon and Maxwell, 2006).[10] Investors with this motivation are clearly interested in accurate measurement of prior social performance, but can rely on social ratings only if they also reliably forecast which firms will exhibit superior social performance in the future, which depends heavily on management decisions made today.

Our second motive, which we term "deontological," applies to investors who do not wish to profit from unethical or heinous actions (Rosen et al., 1991). For example, the Methodist Church's stock market investments have carefully avoided firms involved in alcohol and gambling.[11] Deontological investors care about past performance because they want to ensure that current profits were not earned from prior unethical behavior. They also care deeply about the quality of the current management because they want to avoid future scandals that would taint future profits.

Investors and consumers with consequentialist motives intend their investments and purchases to reward good behavior and to provide an incentive for firms with lagging social performance to improve. They expect their purchases and investments to help responsible firms grow, and reduce market share and raise the cost of capital for environmentally irresponsible firms (Langbein and Posner, 1980; Heal, 2001; Stanley and Herb, 2007). Consequentialist investors rely on accurate information about past performance to ensure that their investment decisions reward and punish the appropriate firms. They also strongly desire forward-looking assessments that will provide incentives for managers to embrace long-term responsibility.

Stakeholders with what we call "expressive" motives use their transactions to express their personal identity to themselves and to others (Williams, 2007). For such stakeholders, "socially responsible investment is an extension of [their] way-of-life…" (Rosen et al., 1991, p. 230).[12]

Expressive investors worry that negative social performance taints companies and, by extension, those who invest in them. The same logic leads expressive investors to be concerned about how today's managerial investments will affect the firm's future reputation.

All four of these motivations described lead social investors to desire transparency about both past social performance and current managerial decisions that will influence future social performance. Social rating agencies such as KLD typically measure both past environmental outcomes and recent management actions that may predict future outcomes (e.g., pollution prevention). In the Appendix, we model how a social rater should optimally weigh the respective information relative to the goals of the rating agency's customers and the information content of each metric. We outline our major results below.

The first result we prove is that a social rater should weigh management quality less heavily when that measure contains little substantial incremental information about future environmental outcomes not contained in history alone (Theorem 1). The intuition behind this result is that if environmental outcomes are similar from year to year, then past harm is already a good measure of current management efforts. Similarly, the social rater should stick largely to historical data if the measure of management quality is extremely noisy.

The second result is that the social rater should weigh management quality more heavily when its customers value transparency about future environmental outcomes (also in Theorem 1). The intuition here is that measures of management quality can yield transparency only with respect to the future; transparency about past performance can be derived directly from measures of past outcomes.

This theorem has two implications. First, in the extreme case in which customers care only about transparency relative to the past, all the social rater needs to deliver is a measure of past

outcomes. Thus, the optimal social rating should have the same ability to predict future harm (as measured by R^2, for example) as the historical data on harm (corollary 1). To the extent the social rater is trying to provide transparency about both past performance and future performance, the optimal social rating should be better able to predict future harm (as measured by a higher R^2, for example) than simply relying on historical harm data (corollary 2).

These corollaries provide testable implications of how well a social rater uses data to satisfy its customers' goals. But because the transparency of social ratings is rarely evaluated, it is unclear whether investors are being well served by social rating agencies. In the following section, we discuss KLD's ratings. In section 4, we outline our methods of testing the ability of KLD's ratings to achieve the two forms of transparency discussed above, namely, summarizing past harm and predicting future harm.

3. Data

Our sample consists of all 588 companies in the United States regulated by the U.S. Environmental Protection Agency whose corporate social performance has been rated by KLD at least once during the period 1991-2003.[13] The firms in our sample are all large and publicly traded, and represent a wide variety of industries (see Table 2). Our sample is an unbalanced panel consisting of data from roughly 350 firms per year, and 3,831 company-year observations for the study period.[14]

3.1 Environmental Ratings

Deckop et al. (2006, p. 334) describe KLD as "the largest multidimensional CSP [corporate social performance] database available to the public." Widely used in studies of corporate social

responsibility (e.g., Graves and Waddock, 1994; Berman et al., 1999), KLD data have recently been referred to as "the de facto [CSP] research standard at the moment" (Waddock, 2003, p. 369). According to KLD, 15 of the top 25 institutional financial managers in the world use its research and more than $10 billion is invested in funds based on its ratings.[15]

KLD employs a proprietary system to evaluate corporations' environmental, social, and governance performance and generate annual company ratings. We obtained data on each of KLD's 14 dichotomous environmental "strength" and "concern" variables. The seven environmental "strength" variables include: *Beneficial products and services*; *Pollution prevention*; *Recycling*; *Clean energy*; *Communications*; *Property, plant, and equipment*; and *Other strength.* The seven environmental "concern" variables include: *Hazardous waste*; *Regulatory problems*; *Ozone-depleting chemicals*; *Substantial emissions*, *Agricultural chemicals*, *Climate change*, and *Other concern*. [16] Detailed descriptions of these ratings are provided in Table 3.

In addition to analyzing the 14 KLD environmental subscores separately, we also aggregate the ratings into *total environmental strengths* and *total environmental concerns*. Finally, we follow the practice, common in the academic literature, of subtracting the concerns from the strengths to arrive at a single *net environmental score* (see, for example, Ruf et al., 2001; Griffin and Mahon, 1997; Johnson and Greening, 1999; Waddock and Graves, 1997; Graves and Waddock, 1994).[17]

3.2 Environmental Performance

Data on companies' environmental performance for the period 1990-2003 were obtained from the Corporate Environmental Profiles Directory (CEPD), which is used by investment professionals, asset managers, and major corporations and financial management companies. Produced by Institutional Shareholder Services (ISS) and the Investor Responsibility Research

Center (IRRC), the CEPD aggregates U.S. EPA environmental data from subsidiary facilities to parent companies in the S&P 1500.

We measure *emissions* as pounds of toxic chemicals reported to the U.S. Environmental Protection Agency's Toxic Release Inventory (TRI) as production waste, transfers, and releases, excluding one-time releases. We measure compliance with environmental regulations as a company's annual *value of penalties assessed* and *number of penalties* associated with violations of nine major federal environmental statutes.[18] To reduce the impact of outliers in our models, we log the pollution and penalty values. To avoid overemphasizing changes from small initial values, we add 1,000 pounds to emissions and $1,000 to penalties prior to taking logs.[19] Similarly, we cap the number of penalties at six per year, the 95th percentile value.

We also consider whether facilities reported to the Emergency Response Notification System *any major spills* of chemicals or oil.[20] Finally, we consider whether facilities incurred any *permit denials* related to the Resource Conservation and Recovery Act (RCRA) or *shut-ins* by the Minerals Management Service (MMS). Permit denial under RCRA occurs when EPA determines that the facility seeking the permit is not properly equipped to treat, store, or dispose of hazardous wastes.[21] Shut-ins refer to instances in which the MMS required an offshore drilling or production facility to cease operations until it remedied unsafe conditions that risked causing oil spills or explosions.[22] Although the Corporate Environmental Profiles Directory provides data on annual numbers of spills, permit denials, and shut-ins, these were zero for the vast majority of firm-year observations (e.g., 98% of firm-years had zero permit denials or shut-ins), and were one for most of the non-zero instances (that is, only a single spill or a single permit denial or shut-in). We consequently coded these as dichotomous variables (zero versus positive).

3.3 Control Variables

We gathered data on company industry and financial metrics—including the 3-digit North American Industry Classification System (NAICS) code, revenues, total assets, net income, total common equity, and total net sales—from Compustat. In our models, we control for company size by including *log revenues* and *log assets.* We add $1,000 to revenues and assets before taking the log to avoid overemphasizing differences from small initial values. We control for industry differences by including *industry dummies* for 3-digit NAICS codes.

4. Methods

All social investors, as explained above, should desire both backward-looking transparency (that is, summing up historical harm) and forward-looking transparency (that is, predicting future harm). In this section, we describe an empirical strategy for testing for these two components of transparency.

4.1 Transparency of Past Performance

As we noted earlier, investors driven by any combination of the four motives described above should want to know whether environmental scores are useful aggregates of historical environmental problems. A good index of harm sums up relevant forms of emissions or regulatory actions. These relationships should be particularly strong for any backward-looking environmental subscores.

We assess the extent to which KLD's environmental ratings render transparent firms' recent environmental performance by estimating the following equation:

$$\textbf{Rating}_{i,t} = F(\beta_1\ \textbf{Environmental Performance}_{i,t-2} + \beta_2\ X_{i,t-2} + v_{i,t}) \qquad (1)$$

We focus on the three disaggregated KLD concern ratings—hazardous waste, regulatory problems, and substantial emissions—whose definitions imply that they reflect firms' historical environmental performance. For example, KLD defines the environmental concern *Regulatory problems* thus: "The company has recently paid substantial fines or civil penalties for violations of air, water, or other environmental regulations, or it has a pattern of regulatory controversies under the Clean Air Act, Clean Water Act, or other major environmental regulations." We also run this model on *total environmental concerns*, which aggregates these and KLD's four other environmental concerns (*Agricultural chemicals*, *Ozone-depleting chemicals, Other concern*, and, since 1999, *Climate change*).

Because KLD is not explicit about how its ratings incorporate prior environmental performance, we use a flexible functional form to represent how they might do so. Specifically, we include in $Environmental\ Performance_{i,t-2}$ several transformations of environmental performance data, namely: log of emissions; log of penalty values; number of violations; number of major spills; number of permit denials or shut-ins; and dummies that indicate whether each performance metric is zero or a positive value. For each of these variables, we include values available to KLD both one and two years before it issued its ratings to accommodate delays by regulators in making data publicly available.[23]

The control variables $X_{i,t-2}$ includes year dummies and two measures of corporate size: log of assets, and log of sales. We run models both with and without industry dummies. We use a Poisson model to estimate the *total environmental concerns* rating (a count variable)[24] and a Probit model to estimate the dichotomous *hazardous waste, regulatory problems,* and *substantial emissions* ratings.

4.2 Transparency Related to Future Performance

We next assess the extent to which ratings provide transparency with respect to the likelihood that current environmental management plans and investments will yield environmental improvement. As before, whatever the combination of motives for social investing that drives them, investors should want to know whether environmental scores are useful predictors of future environmental problems.

Given the high cost of carefully monitoring historical emissions and regulatory actions, most users of social ratings do not track historical environmental performance. For these users, the overall correlation between scores and future harms is particularly relevant. To understand this relationship, we estimate the following equation that predicts environmental emissions and compliance based on KLD ratings for two years earlier:

$$\textbf{Environmental Performance}_{i,t} = F(\beta_1\ \textbf{Rating}_{i,t-2} + \beta_2\ \mathbf{X}_{i,t-2} + \mathbf{v}_{i,t}) \qquad (2)$$

The functional form we estimate (OLS, probit, or negative binomial) depends on the units of the environmental performance metric (continuous, binary, or count data). For each of the five environmental performance metrics, we run separate models in which $Rating_{i,t-2}$ refers to (a) the single *net environmental score,* (b) the subtotals *total environmental strengths* and *total environmental concerns*, and (c) the 14 KLD environmental subscores.[25] The control $X_{i,t-2}$ includes dummies for industry and year, and controls for corporate size (log assets and log sales) following prior empirical models of environmental performance (Balabanis, Phillips, and Lyall, 1998; Sharma, 2000; Johnson and Greening, 1999; Waddock and Graves, 1997). The sign, size, and statistical significance of the coefficient on the ratings (β_1) indicates whether and how the environmental ratings predict environmental performance.

To the extent that the environmental rater is measuring management quality that has not yet affected emissions or penalties, the ratings should also predict future emissions and penalties conditioned on historical performance. These relationships should be particularly strong for any forward-looking environmental subscores. That is, an environmental rating agency can create a CSR metric that predicts future performance merely by examining past performance, even if its ratings add little predictive validity beyond the autocorrelation of emissions and regulatory problems. We are interested in assessing whether KLD ratings have incremental predictive power.

To test this possibility, we estimate the same models as in equation (2), but include as an additional control variable the dependent variable (environmental performance) lagged two years:[26]

$$\textbf{Environmental Performance}_{i,t} = F(\beta'_1\ \textbf{Rating}_{i,t-2} + \beta'_2\ \mathbf{X}_{i,t-2} + \beta'_3\ \textbf{Environmental Performance}_{i,t-2} + v'_{i,t}) \qquad (3)$$

Significant coefficients on the environmental rating (β'_1) in this equation would suggest that KLD's ratings are assessing the effect of current environmental management plans and investments on future environmental performance beyond any autocorrelation in performance.

5. Results

Summary statistics are provided in Table 4. The 14 KLD subscores exhibit fairly low intercorrelation: the mean (median) correlation is 0.01 (-0.01) among the seven KLD strengths subscores and 0.17 (0.15) among the seven KLD concerns subscores. The total KLD strengths and total KLD concerns indices were correlated at 0.25, meaning that firms with more environmental strengths *also* have more environmental concerns.

5.1 Evidence of Transparency with Respect to Past Performance

Do *total KLD environmental concerns* and the three explicitly backward-looking subscores *hazardous waste*, *regulatory problems*, and *substantial emissions* provide transparency about firms' prior environmental performance? The inclusion of various transformations of historical performance introduces multicollinearity. As such, we conduct Wald tests to assess whether the sum of the coefficients on various transformations of each environmental performance metric is a significant predictor of the KLD rating and whether these coefficients jointly differ from zero.

The Wald test of the sum of the coefficients reveals that each of the four KLD metrics is predicted by higher past pollution levels and that each KLD metric except *regulatory problems* is also predicted by more major spills (Table 5). To understand the magnitude of the effects, we hold constant sales and assets and compare a firm with average emissions to a firm with emissions persistently one standard deviation above average.[27] The high-emissions firm had 0.35 higher *total environmental concerns* ratings and were 9.4 percentage points more likely to be rated as having a *hazardous waste* concern, 12.4 percentage points more likely to be rated as having a *regulatory problem* concern, and 10.1 percentage points more likely to be rated as having *substantial emissions* (Table 5, Models 1, 3, 5, 7). Compared to firms with no major spills in the previous two years, those with one major spill in each of those years had a 0.19 higher *total environmental concern* rating and were 6.4% more likely to be rated as having a *substantial emissions* concern (Models 1, 7)**.** In short, KLD's environmental concerns have a statistically significant relationship with past environmental performance.

Statistical significance is necessary, but not sufficient, to show that KLD ratings utilize data optimally to provide transparency about historical environmental performance. An additional test is to identify how often KLD listed *concerns* for firms with unambiguously problematic

environmental performance. Recall that roughly 13% of firm-years recorded concerns about emissions and 19% concerns about regulatory problems (means from Table 4). Presumably, an investor who relies on KLD ratings would like them to identify firms with the highest level of emissions and greatest number of penalties, adjusted for firm size.

We therefore examine the firm-years in the top 3% of emissions per sales and penalties per sales. We chose the somewhat arbitrary cut-off of the top 3% because we perceive that investors interested in transparency with respect to historical emissions and fines would be interested in such outliers (and these cut points are far below the 13-19% who received *concerns* from KLD for these subscores). We calculate averages emissions, fines and sales over the three years prior to KLD issuing its ratings (*t-1* to *t-4*), and then took ratios of these averages. For any firm that appears multiple times in the top 3%, we focus only on the first time it becomes a member of this elite set.

This test looks for false negatives (that is, environmentally harmful firms that KLD did not rate as having the relevant concern).[28] In our dataset, the top 3% threshold corresponds to annual averages over a three year period of 225 tons of toxic chemicals per $100 million in sales and $9600 of annual penalties per $100 million in sales. Of the 32 firms with the top 3% of emissions per sales, KLD rated 53% as having *Substantial emissions*. Of the 35 firms with the top 3% of penalties per sales, KLD rated 49% as having *Regulatory problems*. KLD is, thus, only modestly effective at identifying the most pollution-intensive and non-compliant firms for their size.

We conducted further analysis to assess whether these ratings more accurately identified firms with the absolute highest level of emissions or fines (that is, not scaled by firm size). Our results suggest that they do. Of the 31 firms with the top 3% of total annual emissions (in excess of an annual average of 23 million pounds over a three year period), 81% received KLD's *Substantial emissions* rating. Similarly, of the 31 firms with the top 3% of penalties (in excess of an annual

average of $560,000 over a three year period), KLD rated 68% as having *Regulatory problems.* KLD thus does a better, albeit still quite imperfect, job at identifying firms that objectively have the highest emissions and regulatory concerns when not normalizing by firm size.

5.2 Evidence of Transparency with Respect to Future Performance

We ran a number of models to assess the extent to which KLD's ratings provide transparency about managerial decisions that affect future social performance. We first regressed each environmental performance metric on KLD environmental ratings while controlling for industry, year, and company size (equation 2). We analyzed both aggregated scores and disaggregated subscores. We then repeated the models, conditioning as well on the lagged dependent variable (equation 3).

The single *net environmental score* (KLD strengths minus concerns) was highly statistically significant ($p<0.01$) and of the expected sign (negative) in predicting pollution levels, the value and number of regulatory penalties, and whether firms reported any major oil or chemical spills (Table 6, Models 1–8).

Compared to the sample means, the results of our primary models (without lagged dependent variables) indicate that, on average, having one fewer KLD environmental strength or one more KLD environmental concern doubles expected pollution two years later and increases the value of penalties by 43% ($880)[29], the expected annual number of penalties by 0.12 (from 1.06 to 1.18 penalties per year), and the chance of having any major reportable spills by 11% (from 14.0% to 15.4%) (Table 6; Models 1, 3, 5, and 7). We found no evidence that any of the aggregate KLD scores were associated with subsequent permit denials or shut-ins.

The results reported in Table 7 show that the *total environmental concerns* aggregate KLD score is driving these relationships. That is, additional KLD environmental strengths have no large or statistically significant relationship with our environmental performance metrics, but environmental concerns do. Our results indicate that, compared to the mean outcome levels, a one-unit increase (that is, worsening) a firm's total KLD environmental concerns score is associated with a near tripling of emissions, a 69% increase in the value of penalties ($1,410)[30], a 0.17 increase in the expected number of penalties (from 1.06 to 1.23 per year), and a 19% greater chance of having a major reportable spill (from 14.0% to 16.6%) (Table 7, Models 1, 3, 5, and 7).

To what extent are these results driven by autocorrelation, whereby poor performers continue to perform poorly? Our results indicate that although the magnitude of the *total environmental concern* coefficients decline substantially in the models that include the lagged dependent variable, they remain statistically significant and of the expected sign (Table 7, Models 2, 4, 6, and 8). The decline in the coefficients suggests that a substantial portion of the estimated effect size in our base models is due to autocorrelation. The proportion of the effect size in the base models in Table 7 apparently due to autocorrelation ranges from 19% (spills models) to 73% (pollution models). In the models in which we control for autocorrelation by including the lagged dependent variable, the effects of total KLD environmental concerns remain statistically significant and economically substantial. Compared to the sample means, our results show a one-unit increase in (worsening of) a firm's total environmental concerns score to be associated with 33% higher emissions, 50% higher penalties ($3,064 versus the mean of $2,040), a 0.09 increase in the annual expected number of penalties (from 1.06 to 1.15 per year), and a 15% greater chance of having a major spill (from 14.0% to 16.1%) (Table 7, Models 2, 4, 6, and 8).

We turn now to the results of the models with disaggregated KLD ratings (Table 8). With 14 subscores, 5 environmental performance measures and 2 specifications (with and without historical performance measures), we have many coefficients. Thus, we first summarize the overall pattern of results and then discuss specific subscores. Overall, 66% (87 of 132) of the subscore coefficients across the 10 regressions were of the expected sign in predicting environmental performance (negative for strengths, positive for concerns). Moreover, 65% (34 of 52) of the statistically significant coefficients were of the expected sign in predicting environmental performance.

Several concerns subscores were statistically significant and of the expected sign in predicting multiple environmental performance metrics, regardless of whether the lagged dependent variable was included in the model. For example, firms rated by KLD as having *hazardous waste* concerns subsequently emitted more toxic chemical (as measured by TRI releases) and reported more major spills than other firms of comparable size in the same industry. Firms rated by KLD as having *regulatory problems* subsequently had larger TRI emissions and were assessed more, and more expensive, penalties. Firms rated as having *substantial emissions* were subsequently assessed greater regulatory penalties and reported more major spills. With regard to strengths, firms rated by KLD as having *communications* strengths, in particular, exhibited low penalties two years later relative to firms of comparable size in the same industry.

A few subscores had statistically significant coefficients that were opposite the expected sign. For example, firms KLD rated as having a *climate change* concern were subsequently assessed lower penalties and reported fewer major spills than other firms in the same industry. Similarly, firms rated as having *regulatory problems* subsequently had fewer permit denials and shut-in than other firms in the same industry. But firms rated as having *beneficial products and services*, a KLD strength, had higher TRI emissions, incurred more regulatory penalties (in value and number), and

were more likely to have reported major spills than companies of comparable size in the same industry. Given the high number of coefficients in Table 8, we do not want to over-emphasize a few unexpected results. Nevertheless, these findings merit further investigation in future research.

The KLD subscores that most clearly seek to assess the quality of firms' environmental management practices are two "strengths": *pollution prevention*, and *property, plant and equipment*. The former is awarded when a "company has notably strong pollution prevention programs including both emissions reductions and toxic-use reduction programs," the latter when a "company maintains its property, plant, and equipment with above-average environmental performance for its industry." Accurately measured, these two subscores, above all others, should be associated with improved environmental performance. To measure this association, we ran models that estimated future environmental performance based on these two subscores, controlling for the three concern subscores (*hazardous waste*, *regulatory problems*, and *substantial emissions*) that most clearly aggregate historical performance related to the outcomes we measure.

The results are presented in Table 9. The coefficient on the *pollution prevention* rating indicates that firms with this KLD strength rating incurred lower penalties and had lower TRI emissions and fewer reportable spills, but none of these effects were statistically significant. Our point estimates also suggest that firms rated as having superior *property, plant, and equipment* subsequently had lower TRI emissions and fewer compliance problems and were less likely to report major spills, although only the latter effect was statistically significant.

How much explanatory power?

If KLD ratings accurately measure management actions that lead to future environmental performance, adding KLD ratings to lagged performance data should substantially enhance the explanatory power of the regressions. To assess this, we shift our attention from coefficients on the

KLD ratings to the incremental explanatory power of models that include KLD ratings. Specifically, we examine the magnitude and significance of increases in model fit statistics—adjusted R^2 for OLS models, and McFadden's adjusted R^2 for probit and negative binomial models—across several models. Our baseline model predicts environmental performance based on industry, year, and lagged size controls. The upper left boxed cells in Table 10 present the fit statistics of these models, to which we compare a "KLD model" that adds the 14 lagged KLD subscores to the baseline model (upper right cells in Table 10), a "Lags model" that adds the 1- and 2-year lags of the dependent variable (environmental performance) to the baseline model (lower left cells), and a "Full model" that adds both the KLD subscores and the 1- and 2-year lagged dependent variables (lower right cells).

In predicting emissions, the adjusted R^2 of the full model is 0.832, which is 0.087 larger than the KLD-only model (which omits performance lags) and 0.001 larger than the lags-only model (which omits KLD subscores). Although both increments are statistically significant, adding KLD subscores increases explanatory power only 1% as much as adding lag performance. In predicting the other four performance metrics, including lagged performance consistently provides a much larger boost to explanatory power than including the KLD subscores.

These results are inconsistent with both corollary 1 and corollary 2 in the theoretical model in the Appendix. These corollaries show that if KLD is trying to summarize history or both summarize history and predict future environmental performance, its scores should be at least as useful as history alone in predicting the future. The empirical results in Table 10, however, indicate that KLD has *less* explanatory power than historical performance. Specifically, the R squared values of the "lags models" are consistently greater than the R squared values of the "KLD models". This implies that KLD is not optimally aggregating available data.

But KLD subscores are statistically significant in predicting later emissions and regulatory violations beyond the levels that could otherwise be predicted based on historical performance, company size, and industry. These results are consistent with the hypothesis that KLD raters can identify a small but statistically significant signal of the management practices and organizational capabilities of the firms they rate.

5.3 Robustness Tests

Our main results that predict environmental performance were robust to several alternative specifications. First, we re-estimated our OLS models (with *total environmental strengths* and *total environmental concerns*) using tobit because our continuous dependent variables (log emissions and log penalty values) are left truncated. Second, we re-estimated our models in Table 7 including three additional controls—return on assets, return on equity, and return on sales—because prior research has indicated that profitability influences environmental performance (Johnson and Greening 1999).[31] Third, we ran models that controlled for firm size in the dependent variable rather than as a control variable. Specifically, we regressed *emissions per assets* and *penalty value per assets* on *total environmental concerns*, *total environmental strengths*, and the year and industry dummies.[32] The results of these robustness tests (not shown) were similar to our main results: the coefficients on the *total environmental concerns* variables remained positive and statistically significant, and the coefficients on the *total environmental strengths* variables remained statistically insignificant.

In our main results, we found that in nearly all cases, a significant portion of the relationship between KLD's environmental ratings and subsequent environmental performance remained after we included lagged values of the environmental performance dependent variable, although the

coefficient typically shrank substantially. It is possible that the remaining effect of the KLD ratings might be due to their correlation with lagged values of our *other* environmental performance variables. To test this, we ran models that predicted *emissions* and *penalty value* based on *total environmental strengths* and *total environmental concerns*, controlling for 2-year lags of the other four environmental outcomes (as well as log revenues, log assets, and industry). We also ran models that included additional lags of the dependent variable (specifically, 3- and 4-year lags). The results these robustness tests (not shown) were also similar to our main results: the coefficients on the *total environmental concerns* variables remained positive and statistically significant, and the coefficients on the *total environmental strengths* variables remained statistically insignificant.

6. Discussion

6.1 Summary

Our results assess the transparency of a well-established environmental performance rating system. We have three main results.

1. KLD environmental ratings do a reasonable job of aggregating past environmental performance. For example, higher past pollution levels predict KLD's *total environmental concerns* as well as the three individual KLD concerns we examined: *hazardous waste*, *regulatory problems*, and *substantial emissions*.
2. The single KLD *net environmental score* (environmental strengths ratings minus environmental concerns ratings) and KLD's *total environmental concerns* ratings helped predict future pollution levels, the value and number of subsequent regulatory penalties, and

whether firms eventually reported any major spills. KLD's *total environmental strengths* ratings did not predict subsequent environmental outcomes.

3. The explanatory power of KLD's ratings in predicting future emissions and penalties is far lower than the explanatory power of lagged emissions and penalties. These results imply that KLD is not optimally aggregating historical data, regardless of how it weights historical performance versus management quality.

6.2 Limitations and Future Research

This paper is to our knowledge the first to examine the transparency of social or environmental ratings. As such, it has a number of limitations, each of which points the way to future research.

We found little evidence that KLD's environmental strengths predicted any of the environmental outcomes we analyzed. This result might be due in part to our focus on pollution and regulatory violations, environmental "bads" rather than environmental "goods." That is, our results do not examine the ability of KLD ratings to predict significant environmental successes through new products or other means (as measured by KLD's *beneficial products & services* subscore, for example). But it is surprising that we found no evidence that *pollution prevention*, a KLD strength, predicted either pollution or regulatory violations. Thus, future research should examine the predictive validity of CSR ratings on positive environmental outcomes such as developing innovative green products. In addition, the ratings of other social rating agencies should be examined.

Furthermore, the predictive component of transparency suffers from the lack of an absolute standard. Conditioning on past values of emissions (together with our standard controls), the 14 KLD environmental subscores added only 0.001 to the adjusted R^2 in predicting future emissions. Is

this effect large or small? For stakeholders, the answer must depend on the cost of the KLD data, whether they have access to data on lagged emissions (if not, the incremental adjusted R^2 rises to 0.087), whether the modest increase in predictive power of emissions is useful, and whether emissions is a large portion of their social evaluation function. Future research should explore the value functions of socially conscious investors, customers, and others to determine the cost-effectiveness of KLD's or others' ratings.

The link between corporate social performance and corporate financial performance should be reexamined in light of our findings. After all, if social ratings are not providing adequate transparency, stakeholders may be responding more to measurement error than actual corporate social responsibility. Future research should examine how the holdings of socially responsible funds change as stakeholders are provided with more transparency about corporate social performance. Interestingly, there may be heterogeneity among stakeholders in how they respond to higher quality information.

It might also be that the act of being rated by KLD has an impact on corporate social performance. By comparing firms that have been rated by KLD to similar firms that have not, researchers could isolate any positive or negative impact of being rated (Chatterji and Toffel, 2007).[33]

Due to data limitations, we have studied the emissions of the U.S.-based facilities owned by each company. In fact, many of the firms in our sample have foreign operations and almost all operate complex supply chains. Future research should thus examine whether and how the accuracy of KLD's ratings is affected by differences in environmental performance in U.S. and foreign subsidiaries and differences within the company versus among suppliers. For example, if a firm allocated environmentally damaging activities to overseas subsidiaries and KLD recognized this,

the firm's KLD rating would decline but our U.S.-based performance metrics would underestimate the relationship between KLD ratings and environmental performance. Conversely, if KLD correctly measured low emissions from a company but ignored its environmentally destructive supply chain, our estimates might over-estimate the relationship between KLD ratings and environmental performance.

We have examined the validity of environmental ratings only in terms of summarizing the past and predicting the future. Others have examined convergent validity (that is, whether different social raters provide similar ratings; see Chatterji and Levine, 2006; Sharfman, 1996). Additional research should examine other forms of validity (e.g., using fair procedures or having high face validity) and additional dimensions of social performance (e.g., treatment of workers, communities, and customers).

6.3 Conclusions

KLD expends substantial resources attempting to measure the quality of companies' environmental management systems. Our results suggest that this is difficult to do well.

Our results are consistent with a large literature that finds low validity of management system measurements. In many surveys, for example, union and management disagree on how often they meet (Eaton, 1994) and whether work has become more intense (Green, 2004). Often, employees do not even agree on the human resource management practices that prevail in their workplace. Such large variation in employees' perceptions of their own firms' management practices is suggestive of the substantial challenge rating agencies face. An even larger literature emphasizes the difficulty humans have deciphering one another's intentions (see, for example, Ross, 1977).

The substantially higher predictive ability of simple autocorrelations over sophisticated judgment models that we find here is consistent with substantial past research. That is, in domains from college admissions to medical diagnoses, simple rules that summarize objective information almost always outperform decisions based on the same information plus qualitative information (that should be useful) when the qualitative information is accompanied by subjective judgment (that can add noise).[34] Furthermore, as discussed above, KLD ratings do a good job of identifying firms with the worst absolute historical environmental performance (not normalized by size), which suggests the ratings might be useful to activists and others who seek to target such firms (Baron and Diermeier, forthcoming). But KLD ratings do a much poorer job identifying the worst performing firms adjusted by size, suggesting that investors interested in firms' eco-efficiency ought to interpret KLD ratings with caution.

In short, given the limited validity of the measurement of management systems, it would seem that the validity of KLD's ratings could be improved if substantially more weight were put on historical environmental performance. Moreover, there is no reason to summarize individual environmental subscores such as emissions or regulatory problems as a one or zero indicator variable. The rich data on environmental outcomes that we analyze is also available to KLD, which could summarize environmental strengths and concerns much more accurately if it used a continuous or multi-level indicator. A simplified scale might be more relevant to social and environmental metrics that lack supporting quantitative data.

In sum, an important challenge for researchers and environmental stakeholders alike is to identify valid measures of the quality of environmental management systems. We hope that the somewhat discouraging results we present here will motivate future research. All stakeholders—from investors, employees, and boards of directors to customers, regulators, and activists—have an

interest in the social and environmental performance of enterprises. These topics are too important to rely on metrics that are untested or that, when tested, do not provide transparency.

7. References

Bagnoli, M. and S.G. Watts, 2003, "Selling to Socially Responsible Consumers: Competition and the Private Provision of Public Goods," *Journal of Economics and Management Strategy*, 12(3), 419–445.

Balabanis, G., H.C. Phillips, and J. Lyall, 1998, "Corporate Social Responsibility and Economic Performance in the Top British Companies: Are They Linked?" *European Business Review*, 98(1), 25–44.

Baron, D.P., 2001, "Private Politics, Corporate Social Responsibility, and Integrated Strategy," *Journal of Economics and Management Strategy*, 10(1), 7–45.

Baron, D.P. and D. Diermeier, forthcoming, "Strategic Activism and Nonmarket Strategy," *Journal of Economics and Management Strategy*.

Berman, S.L., A.C. Wicks, S. Kotha, and T.M. Jones, 1999, "Does Stakeholder Orientation Matter? The Relationship between Stakeholder Management Models and Firm Financial Performance," *Academy of Management Journal*, 42, 488–506.

Cantor, R., 2004, "An Introduction to Recent Research on Credit Ratings," *Journal of Banking and Finance*, 28, 2565–2573.

Chatterji, A.K. and D.I. Levine, 2006, "Irresponsibility in Measuring Corporate Responsibility," Working Paper, University of California at Berkeley.

Chatterji, A.K., and M.W. Toffel, 2007, "Shamed and Able: How Firms Respond to Information Disclosure," Working Paper, Harvard Business School.

Dawes, R.M., D. Faust, and P.E. Meehl, 1989, "Clinical versus Actuarial Judgment," *Science*, 243, 1668–1674.

Deckop, J.R., K.K. Merriman, and S. Gupta, 2006, "The Effects of CEO Pay Structure on Corporate Social Performance," *Journal of Management*, 32(3), 329–342.

Eaton, A.E., 1994, "The Survival of Employee Participation Programs in Unionized Settings," *Industrial and Labor Relations Review*, 47(3), 371–389.

Entine, J., 2006, "Is Corporate Social Responsibility Serious Business?" *Ethical Corporation* (May).

Graves, S. and S.A. Waddock, 1994, "Institutional Owners and Corporate Social Performance," *Academy of Management Journal*, 37(4), 1034–1046.

Green, F., 2004, "Why Has Work Effort Become More Intense?" *Industrial Relations*, 43(4), 709–741.

Griffin, J.J. and J.F. Mahon, 1997, "The Corporate Social Performance and Corporate Financial Performance Debate: Twenty-Five Years of Incomparable Research," *Business and Society,* 36(1), 5–31.

Güttler, A., 2005, "Using a Bootstrap Approach to Rate the Raters," *Financial Markets and Portfolio Management*, 19, 277–295.

Hart, S., 1995, "A Natural-Resource-Based View of the Firm," *Academy of Management Review*, 20, 986–1014.

Heal, Geoffrey, 2001, "Survey-Mastering Investment: The Bottom Line to a Social Conscience," *The Financial Times*, July 2.

IRRC, 2002, *IRRC's Corporate Environmental Profile Database: Purpose, Methodology and Limitations*, Washington, D.C.: Investor Responsibility Research Center.

Johnson, R.A. and D.W. Greening, 1999, "The Effects of Corporate Governance and Institutional Ownership Types on Corporate Social Performance," *Academy of Management Journal,* 42(5), 564–576.

Keele, L. and N.J. Kelly, 2006, "Dynamic Models for Dynamic Theories: The Ins and Outs of Lagged Dependent Variables," *Political Analysis* 14(2):186-205.

Krämer, W. and A. Güttler, 2006, "On Comparing the Accuracy of Default Predictions in the Rating Industry," Working Paper, Business and Social Statistics Department, University Dortmund [http://econpapers.repec.org/paper/dorwpaper/2.htm].

Langbein, J.H. and R.A. Posner, 1980, "Social Investing and the Law of Trusts," *Michigan Law Review*, 79(1), 72–112.

Lev, Baruch, Christine Petrovits, and Suresh Radhakrishnan, 2006, "Is Doing Good Good for You? Yes, Charitable Contributions Enhance Revenue Growth," Working Paper, New York University Stern School of Business.

Lyon, T.P. and J.W. Maxwell, 2006, "Greenwash: Corporate Environmental Disclosure under Threat of Audit," Working Paper, University of Michigan, Stephen M. Ross School of Business.

Margolis, J. and J. Walsh, 2003, "Misery Loves Company: Rethinking Social Initiatives by Business," *Administrative Sciences Quarterly*, 48, 268–305.

Margolis, J., H.A. Elfenbein and J. Walsh, 2007, "Does It Pay To Be Good? A Meta-Analysis and Redirection of Research on the Relationship Between Corporate Social and Financial Performance," mimeo.

Mattingly, J.E. and S.L. Berman, 2006, "Measurement of Corporate Social Action: Discovering Taxonomy in Kinder, Lydenberg, Domini Ratings Data," *Business and Society*, 45(1), 20–46.

Maxwell, J.W., T.P. Lyon, and S.C. Hackett, 2000, "Self Regulation and Social Welfare: The Political Economy of Corporate Environmentalism," *Journal of Law and Economics*, 43 (2), 583–618.

Mora, N., 2006, "Sovereign Credit Ratings: Guilty Beyond Reasonable Doubt?" *Journal of Banking & Finance*, 30, 2041–2062.

Orlitzky, M., F.L. Schmidt, and S.L. Rynes, 2003, "Corporate Social and Financial Performance: A Meta-Analysis," *Organization Studies,* 24, 403–441.

Raftery, A.E., 1995, "Bayesian model selection in social research," in P. V. Marsden, ed., *Sociological Methodology*, San Francisco: Jossey-Bass, 111–163.

Rosen, B.N., D.M. Sandler, and D. Shani, 1991, "Social Issues and Socially Responsible Investment Behavior: A Preliminary Empirical Investigation," *Journal of Consumer Affairs*, 25 (2), 221–234.

Ross, L., 1977, "The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process," in L. Berkowitz, ed., *Advances in Experimental Social Psychology*, 10, New York: Academic Press, 173–220.

Ruf, B.M., K. Muralidhar, R.M. Brown, J.J. Janney, and K. Paul, 2001, "An Empirical Investigation of the Relationship between Change in Corporate Social Performance and Financial Performance: A Stakeholder Theory Perspective," *Journal of Business Ethics*, 32, 143–156.

Russo, M.V. and P.A. Fouts, 1997, "A Resource-Based Perspective on Corporate Environmental Performance and Profitability," *Academy of Management Journal*, 40(3), 534–559.

Sharfman, M.P., 1996, "The Construct Validity of the Kinder, Lydenberg & Domini Social Performance Ratings Data," *Journal of Business Ethics*, 15(3), 287–296.

Sharma, S., 2000, "Managerial Interpretations and Organizational Context as Predictors of Corporate Choice of Environmental Strategy," *Academy of Management Journal*, 43(4), 681–694.

Stanley, D.J. and C.R. Herb, 2007, "The Moral and Financial Conflict of Socially Responsible Investing," *Graziadio Business Report*, 10(1).

Sy, A., 2004, "Rating the Rating Agencies: Anticipating Currency Crises or Debt Crises?" *Journal of Banking and Finance*, 28, 2845–2867.

Waddock, S., 2003, "Myths and Realities of Social Investing," *Organization and Environment*, 16, 369–380.

Waddock, S. A. and S. Graves, 1997, "The Corporate Social Performance–Financial Performance Link," *Strategic Management Journal*, 18, 303–317.

Williams, G., 2007, "Some Determinants of the Socially Responsible Investment Decision: A Cross-Country Study," *Journal of Behavioral Finance*, 8(1), 43-57.

Table 1: Motives of Socially Responsible Investors

Investor motivation to use environmental ratings	Importance of social ratings incorporating…	
	…firms' prior record	**…management quality**
Financial investors believe that superior environmental performance leads to superior financial performance.	Identify firms' prior environmental problems, which indicate high expected liability and future penalties.	Better quality of environmental management is predictive of future profitability via less waste, lower compliance penalties, and better reputation with stakeholders and is indicative of better management quality more broadly.
Deontological investors seek to avoid investments in companies that act irresponsibly towards the environment because they consider it unethical to earn profits from irresponsible companies.	Identify firms' prior environmental problems, which taint current profits.	Identify responsible management practices to avoid tainted future profits.
Consequentialist investors seek to direct their funds to raise the cost of capital for misbehaving firms and lower it for socially responsible firms.	Identify firms' environmental records to punish or reward firms, as appropriate.	Provide incentives for current investment in long-term environmental performance by rewarding responsible management decisions today.
Expressive investors base their social identity in part on their investments and associations with good causes, and thus seek to invest in companies widely perceived to be "environmentally responsible."	Avoid firms with prior environmental problems to avoid social stigma; seek firms with favorable environmental records to bolster social status.	Seek high quality environmental management to avoid future social stigmas and to enhance future social status.

Table 2. Sample Distribution by Industry

NAICS Code	Industry Description	Frequency
11	Agriculture, forestry, fishing and hunting	2
21	Mining	34
22	Utilities	58
23	Construction	10
311	Food manufacturing	27
312	Beverage and tobacco product manufacturing	9
314	Textile product mills	2
315	Apparel manufacturing	6
316	Leather and allied product manufacturing	5
321	Wood product manufacturing	5
322	Paper manufacturing	22
323	Printing and related support activities	7
324	Petroleum and coal products manufacturing	14
325	Chemical manufacturing	73
326	Plastics and rubber products manufacturing	14
327	Nonmetallic mineral product manufacturing	9
331	Primary metal manufacturing	24
332	Fabricated metal product manufacturing	27
333	Machinery manufacturing	50
334	Computer and electronic product manufacturing	102
335	Electrical equipment, appliance, and component manufacturing	23
336	Transportation equipment manufacturing	49
337	Furniture and related product manufacturing	10
339	Miscellaneous manufacturing	24
42	Wholesale trade	23
44-45	Retail trade	28
48-49	Transportation and warehousing	22
51	Information	22
52	Finance and insurance	29
53	Real estate and rental and leasing	5
54	Professional, scientific, and technical services	7
56	Administrative and support and waste management and remediation services	7
62	Health care and social assistance	4
71	Arts, entertainment, and recreation	1
72	Accommodation and food services	4
81	Other services (except public administration)	2
	Not available	5
	Total:	765

Table 3.Description of KLD Environmental Ratings (as of 2006)

KLD environmental strengths

1. **Beneficial products and services**. The company derives substantial revenues from innovative remediation products, environmental services, or products that promote the efficient use of energy, or it has developed innovative products with environmental benefits. (The term "environmental service" does not include services with questionable environmental effects such as landfills, incinerators, waste-to-energy plants, and deep injection wells.)

2. **Pollution prevention**. The company has notably strong pollution prevention programs including both emissions reductions and toxic-use reduction programs.

3. **Recycling**. The company either is a substantial user of recycled materials as raw materials in its manufacturing processes,or a major factor in the recycling industry.

4. **Clean energy** (previously called Alternative fuels). The company has taken significant measures to reduce its impact on climate change and air pollution through use of renewable energy and clean fuels or through energy efficiency. The company has demonstrated a commitment to promoting climate-friendly policies and practices outside its own operations.

5. **Communications**. The company is a signatory to the CERES Principles, publishes a notably substantive environmental report, or has notably effective internal communications systems in place for environmental best practices. KLD began assigning strengths for this issue in 1996.[35]

6. **Property, plant, and equipment**. The company maintains its property, plant, and equipment with above-average environmental performance for its industry. KLD has not assigned strengths for this issue since 1995.

7. **Other strength**. The company has demonstrated a superior commitment to management systems, voluntary programs, or other environmentally proactive activities.

KLD environmental concerns

1. **Hazardous waste.** The company's liabilities for hazardous waste sites exceed $50 million, or the company has recently paid substantial fines or civil penalties for waste management violations.

2. **Regulatory problems**. The company has recently paid substantial fines or civil penalties for violations of air, water, or other environmental regulations, or it has a pattern of regulatory controversies under the Clean Air Act, Clean Water Act, or other major environmental regulations.

3. **Ozone-depleting chemicals**. The company is among the top manufacturers of ozone-depleting chemicals such as HCFCs, methyl chloroform, methylene chloride, or bromines.

4. **Substantial emissions**. The company's legal emissions of toxic chemicals (as defined by and reported to the EPA) from individual plants into the air and water are among the highest of the companies followed by KLD.

5. **Agricultural chemicals**. The company is a substantial producer of agricultural chemicals, i.e., pesticides or chemical fertilizers.

6. **Climate change**. The company derives substantial revenues from the sale of coal or oil and its derivative fuel products, or the company derives substantial revenues indirectly from the combustion of coal or oil and its derivative fuel products. Such companies include electric utilities, transportation companies with fleets of vehicles, auto and truck manufacturers, and other transportation equipment companies.

7. **Other concern**. The company has been involved in an environmental controversy that is not covered by other KLD ratings.

Source: KLD Ratings Methodology: http://www.kld.com/research/data/KLD_Ratings_Methodology.pdf.

Table 4. Descriptive Statistics

Panel A: Variable Definitions and Summary Statistics

		Obs	Mean	Std. Dev.	Min	Max
Environmental outcomes	Pounds of toxic chemical emissions	4,237	5.1 million	29.3 million	0	743 million
	Log (+1,000) pounds of toxic chemical emissions	4,237	11.55	3.36	6.91	20.43
	Regulatory penalty value	4,237	69,302	568,469	0	25.7 million
	Log (+1,000) regulatory penalty value	4,237	8.02	1.90	6.91	17.06
	Number of regulatory penalties	4,237	1.06	1.93	0	7
	Any major chemical or oil spills	4,237	0.14	0.34	0	1
	Any permit denials or shut-ins	4,237	0.03	0.17	0	1
KLD aggregated scores	Net KLD environmental score (strengths minus concerns)	3,742	-0.29	1.06	-5	3
	Total KLD environmental strengths	3,742	0.34	0.59	0	3
	Total KLD environmental concerns	3,743	0.64	1.04	0	6
KLD strength subscores	Beneficial products and services	3,743	0.07	0.25	0	1
	Pollution prevention	3,743	0.07	0.26	0	1
	Recycling	3,743	0.05	0.22	0	1
	Alternative fuels	3,743	0.06	0.25	0	1
	Communications (added in 1996)	2,612	0.06	0.24	0	1
	Property, plant & equipment (until 1995)	1,131	0.00	0.06	0	1
	Other strength	3,743	0.04	0.20	0	1
KLD concern subscores	Hazardous waste	3,743	0.20	0.40	0	1
	Regulatory problems	3,743	0.19	0.39	0	1
	Ozone depleting chemicals	3,743	0.01	0.09	0	1
	Substantial emissions	3,743	0.13	0.33	0	1
	Agricultural chemicals	3,743	0.02	0.15	0	1
	Climate change (added in 1999)	1,882	0.13	0.34	0	1
	Other concern	3,743	0.03	0.17	0	1
Controls	Log sales (+ 1,000)	4,237	22.20	1.52	15.89	28.03
	Log assets (+ 1,000)	4,237	22.04	1.35	16.56	26.38
	Return on assets	4,237	0.05	0.07	-0.36	0.24
	Return on sales	4,237	0.05	0.10	-0.85	0.35
	Return on equity	4,236	0.12	0.29	-1.46	1.62

Panel B: Pair-Wise Correlations

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	Log (pounds transferred & released + 1000)	1.00																										
2	Log (regulatory penalty value + $1000)	0.29	1.00																									
3	Number of regulatory penalties	0.29	0.78	1.00																								
4	Any major chemical or oil spills	0.23	0.39	0.42	1.00																							
5	Any permit denials or shut-ins	-0.02	0.14	0.17	0.19	1.00																						
6	Net KLD environmental score	-0.37	-0.31	-0.34	-0.32	-0.05	1.00																					
7	Total KLD environmental strengths	0.19	0.15	0.15	0.15	0.12	0.31	1.00																				
8	Total KLD environmental concerns	0.48	0.40	0.43	0.41	0.12	-0.84	0.25	1.00																			
9	KLD Strength: Beneficial products & services	0.05	-0.01	-0.01	0.00	-0.03	0.27	0.42	-0.04	1.00																		
10	KLD Strength: Pollution prevention	0.14	0.09	0.06	0.05	0.03	0.12	0.50	0.17	0.02	1.00																	
11	KLD Strength: Recycling	0.17	0.15	0.14	0.06	0.00	0.10	0.36	0.10	-0.04	0.01	1.00																
12	KLD Strength: Alternative fuels	-0.03	0.05	0.09	0.14	0.33	0.07	0.40	0.15	0.00	-0.04	-0.03	1.00															
13	KLD Strength: Communications	0.15	0.09	0.09	0.13	-0.02	0.17	0.57	0.17	0.05	0.22	0.06	0.05	1.00														
14	KLD Strength: Property, plant & equipment	-0.08	0.02	0.00	0.00	-0.01	0.08	0.08	-0.04	-0.01	-0.02	-0.02	-0.02	.	1.00													
15	KLD Strength: Other strength	0.04	0.06	0.06	0.05	-0.04	0.08	0.35	0.12	-0.02	-0.01	-0.01	0.00	0.10	-0.01	1.00												
16	KLD Concern: Hazardous waste	0.33	0.35	0.38	0.35	0.12	-0.60	0.20	0.73	-0.03	0.16	0.05	0.13	0.15	-0.03	0.09	1.00											
17	KLD Concern: Regulatory problems	0.32	0.35	0.35	0.34	0.08	-0.63	0.19	0.75	-0.01	0.14	0.13	0.09	0.11	-0.03	0.05	0.44	1.00										
18	KLD Concern: Ozone depleting chemicals	0.15	0.16	0.17	0.14	0.00	-0.24	0.12	0.31	0.01	0.06	-0.01	0.04	0.02	-0.01	0.20	0.15	0.12	1.00									
19	KLD Concern: Substantial emissions	0.42	0.28	0.28	0.29	0.04	-0.57	0.18	0.68	-0.03	0.12	0.14	0.08	0.13	-0.02	0.06	0.34	0.37	0.18	1.00								
20	KLD Concern: Agricultural chemicals	0.17	0.18	0.20	0.17	0.00	-0.32	0.10	0.38	-0.01	0.09	-0.03	0.03	0.05	-0.01	0.13	0.21	0.19	0.43	0.17	1.00							
21	KLD Concern: Climate change	0.31	0.09	0.16	0.16	0.12	-0.48	0.08	0.54	-0.06	-0.04	-0.05	0.20	0.12	.	0.02	0.14	0.24	-0.03	0.26	-0.05	1.00						
22	KLD Concern: Other concern	0.10	0.09	0.11	0.07	0.07	-0.26	0.05	0.29	-0.04	0.04	0.05	0.03	-0.01	-0.01	0.05	0.08	0.11	0.02	0.09	0.01	0.08	1.00					
23	Log sales (+ $1,000)	0.22	0.27	0.25	0.25	0.09	-0.32	0.20	0.44	-0.08	0.16	0.02	0.11	0.22	0.06	0.09	0.38	0.31	0.10	0.30	0.09	0.22	0.12	1.00				
24	Log assets (+ $1,000)	0.27	0.26	0.23	0.23	0.03	-0.30	0.18	0.41	-0.05	0.19	0.02	0.01	0.22	0.07	0.12	0.37	0.30	0.11	0.26	0.11	0.16	0.13	0.91	1.00			
25	Return on assets	-0.04	-0.03	-0.05	-0.06	-0.03	0.08	-0.06	-0.11	-0.01	0.03	-0.03	-0.09	-0.04	-0.05	0.00	-0.11	-0.07	0.00	-0.10	0.02	-0.06	0.01	-0.09	0.00	1.00		
26	Return on sales	-0.07	0.02	0.00	0.01	0.04	0.03	-0.04	-0.06	-0.04	0.01	-0.04	-0.01	-0.01	-0.06	-0.02	-0.04	-0.04	0.01	-0.06	0.02	-0.01	0.01	0.06	0.02	0.83	1.00	
27	Return on equity	0.00	-0.01	-0.03	-0.03	-0.02	0.02	-0.02	-0.03	0.00	0.01	-0.02	-0.05	-0.01	-0.06	0.00	-0.04	0.00	0.01	-0.05	0.02	-0.03	0.01	0.03	0.08	0.57	0.48	1.00

Table 5. What Determines Environmental Ratings? Predicting KLD Environmental Scores

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total KLD environmental concerns		KLD Concern: Hazardous waste		KLD Concern: Regulatory problems		KLD Concern: Substantial emissions	
	Poisson		Probit		Probit		Probit	
No TRI emissions, lagged 3 years	0.420	0.151	-0.009	-0.061	0.056	0.041	0.105	0.141
	[0.176]**	[0.106]	[0.065]	[0.067]	[0.070]	[0.084]	[0.111]	[0.143]
No TRI emissions, lagged 4 years	0.529	0.246	0.200	0.186	0.200	0.189	0.076	-0.021
	[0.162]***	[0.093]***	[0.089]**	[0.119]	[0.074]***	[0.084]**	[0.070]	[0.025]
Log (+1,000) TRI emissions, lagged 3 years	0.069	0.048	0.006	0.011	0.018	0.023	0.019	0.022
	[0.011]***	[0.009]***	[0.008]	[0.009]	[0.007]**	[0.009]***	[0.005]***	[0.006]***
Log (+1,000) TRI emissions, lagged 4 years	0.036	0.030	0.022	0.019	0.019	0.018	0.011	0.006
	[0.009]***	[0.007]***	[0.008]***	[0.010]*	[0.007]***	[0.008]**	[0.004]***	[0.004]*
No penalty amount, lagged 1 year	-0.179	-0.071	-0.074	0.001	-0.064	-0.030	-0.022	-0.015
	[0.077]**	[0.056]	[0.064]	[0.063]	[0.059]	[0.060]	[0.027]	[0.028]
No penalty amount, lagged 2 years	-0.071	-0.008	-0.042	0.008	0.059	0.088	-0.005	0.000
	[0.075]	[0.058]	[0.055]	[0.059]	[0.057]	[0.058]	[0.021]	[0.024]
Log penalty amount, lagged 1 year	0.009	0.008	0.001	0.003	0.013	0.009	0.001	0.001
	[0.008]	[0.006]	[0.007]	[0.007]	[0.007]*	[0.008]	[0.002]	[0.003]
Log penalty amount, lagged 2 years	0.017	0.016	0.002	0.001	0.025	0.023	0.003	0.003
	[0.007]**	[0.006]***	[0.007]	[0.007]	[0.007]***	[0.007]***	[0.002]	[0.003]
Number of fines, lagged 1 year	0.009	-0.001	0.010	-0.001	0.004	-0.005	-0.004	-0.009
	[0.028]	[0.021]	[0.026]	[0.026]	[0.023]	[0.023]	[0.009]	[0.011]
Number of fines, lagged 2 years	-0.009	-0.007	0.010	0.005	-0.018	-0.026	-0.002	-0.006
	[0.026]	[0.021]	[0.022]	[0.022]	[0.025]	[0.026]	[0.008]	[0.009]
Log number of fines, lagged 1 year	-0.065	-0.017	-0.026	0.030	-0.031	0.012	-0.001	0.010
	[0.125]	[0.095]	[0.106]	[0.112]	[0.097]	[0.099]	[0.039]	[0.045]
Log number of fines, lagged 2 years	0.051	0.045	-0.003	0.032	0.091	0.124	0.002	0.013
	[0.120]	[0.096]	[0.095]	[0.099]	[0.105]	[0.108]	[0.033]	[0.039]
Any major spills, 1 year ago	0.100	0.062	0.055	0.034	0.023	0.009	0.038	0.040
	[0.033]***	[0.025]**	[0.033]*	[0.032]	[0.031]	[0.031]	[0.017]**	[0.020]**
Any major spills, 2 years ago	0.086	0.050	0.045	0.028	-0.007	-0.032	0.032	0.030
	[0.033]***	[0.026]*	[0.036]	[0.035]	[0.029]	[0.027]	[0.016]**	[0.016]*
Number of major spills, 1 year ago	0.002	-0.007	0.015	0.006	0.017	0.008	-0.005	-0.004
	[0.007]	[0.007]	[0.012]	[0.012]	[0.011]	[0.012]	[0.003]	[0.004]
Number of major spills, 2 years ago	0.002	-0.002	0.027	0.016	0.018	0.013	-0.001	0.001
	[0.007]	[0.005]	[0.013]**	[0.012]	[0.012]	[0.011]	[0.003]	[0.003]

Any permit denials or shut-ins, 1 year ago	0.132	0.073	0.120	0.146	0.030	-0.025	-0.014	-0.017
	[0.111]	[0.065]	[0.151]	[0.156]	[0.148]	[0.121]	[0.016]	[0.017]
Any permit denials or shut-ins, 2 years ago	0.209	0.081	0.046	0.039	0.227	0.160	0.130	0.137
	[0.090]**	[0.064]	[0.156]	[0.143]	[0.154]	[0.129]	[0.084]	[0.084]
Number of permit denials and shut-ins,	-0.071	-0.069	-0.007	-0.031	-0.033	-0.023	-0.001	0.001
1 year ago	[0.027]***	[0.024]***	[0.040]	[0.039]	[0.048]	[0.062]	[0.010]	[0.014]
Number of permit denials and shut-ins,	-0.057	-0.055	0.002	-0.024	-0.042	-0.048	-0.018	-0.020
2 years ago	[0.032]*	[0.027]**	[0.055]	[0.051]	[0.040]	[0.038]	[0.011]	[0.011]*
Log sales, 1 year ago	-0.032	0.015	0.005	0.033	0.001	0.000	-0.025	-0.020
	[0.039]	[0.031]	[0.032]	[0.034]	[0.031]	[0.032]	[0.012]**	[0.014]
Log sales, 2 years ago	0.012	0.007	0.020	0.025	0.023	0.023	-0.015	-0.015
	[0.038]	[0.031]	[0.030]	[0.033]	[0.034]	[0.039]	[0.011]	[0.014]
Log assets, 1 year ago	0.031	0.042	0.064	0.065	-0.047	-0.065	0.012	0.023
	[0.054]	[0.040]	[0.033]*	[0.039]*	[0.042]	[0.045]	[0.013]	[0.017]
Log assets, 2 years ago	0.092	0.043	-0.001	-0.003	0.072	0.101	0.036	0.030
	[0.060]	[0.047]	[0.037]	[0.044]	[0.045]	[0.050]**	[0.016]**	[0.018]*
Year dummies (1993-2003)	Included	Included	Included	Included	Included	Included	Included	Included
Industry dummies (3-digit NAICS Codes)		Included		Included		Included		Included
Observations (firm-years)	2615	2615	2754	2396	2754	2521	2754	2390
Firms	536	536	570	485	570	514	570	483
McFadden Adjusted R^2	0.289	0.319	0.341	0.385	0.281	0.281	0.431	0.434
Wald χ^2 tests:								
Emissions coefficients jointly = 0?	100.47***	94.13***	22.26***	22.58***	29.06***	28.45***	126.24***	139.12***
Penalty value coefficients jointly = 0?	12.23**	8.94*	1.72	0.25	19.43***	14.44***	3.66	2.36
Penalty number coefficients jointly = 0?	2.47	2.59	2.43	3.77	3.10	2.94	4.30	7.67
Spill coefficients jointly = 0?	19.07***	7.20	18.80***	5.20	9.36*	4.04	13.39**	11.46**
Denial/shut-in coefficients jointly = 0?	11.88**	10.27**	2.72	2.80	6.04	3.69	12**	11.4**
Sum of emissions coefficients = 0?	40.41***	14.86***	5.83**	1.36	11.27***	6.36**	9.46***	1.88
Sum of penalty value coefficients = 0?	2.68	0.29	1.14	0.02	0.12	0.75	0.31	0.05
Sum of penalty number coefficients = 0?	0.01	0.02	0.00	0.20	0.11	0.57	0.01	0.02
Sum of spill coefficients = 0?	13.85***	6.37**	9.47***	2.9*	1.45	0.01	13.28***	10.99***
Sum of denial/shut-in coefficients = 0?	3.00*	0.10	1.00	0.45	1.12	0.08	0.89	0.74

This table reports marginal effects. Brackets contain standard errors clustered by firm. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 6. Results of Models with a Single KLD Environmental Rating

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	Log pounds emissions		Log regulatory penalty value		Number of regulatory penalties		Any major spills		Any permit denials or shut-ins	
Functional form	OLS		OLS		Negative binomial		Probit		Probit	
Net KLD environmental score	-0.716	-0.189	-0.255	-0.189	-0.117	-0.060	-0.015	-0.012	0.004	0.002
(strengths-concerns)	[0.086]***	[0.042]***	[0.055]***	[0.043]***	[0.023]***	[0.017]***	[0.007]**	[0.006]*	[0.005]	[0.004]
Log assets, 2 years ago (+1,000)	0.250	0.169	0.384	0.307	0.134	0.123	0.064	0.053	0.019	0.007
	[0.240]	[0.112]	[0.130]***	[0.111]***	[0.072]*	[0.056]**	[0.014]***	[0.013]***	[0.007]***	[0.006]
Log sales, 2 years ago (+1,000)	0.516	0.098	-0.088	-0.089	0.040	-0.008	-0.022	-0.016	-0.014	-0.003
	[0.253]**	[0.119]	[0.140]	[0.120]	[0.076]	[0.060]	[0.015]	[0.014]	[0.008]*	[0.007]
Lagged dependent variable		Included		Included		Included		Included		Included
Year dummies	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
Industry dummies (3-digit NAICS Codes)	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
Observations (firm-years)	3742	3304	3137	2873	3100	2857	2762	2587	935	842
Firms	744	644	534	451	516	445	431	382	196	165
Adjusted R^2 †	0.474	0.737	0.311	0.372	0.129	0.168	0.300	0.327	0.410	0.503

All independent variables and controls are 2-year lags. For Negative Binomial and Probit models, this table reports marginal effects, not coefficients. Brackets contain standard errors clustered by firm. * significant at 10%; ** significant at 5%; *** significant at 1%. The samples in the Probit models are smaller because these models drop observations from industries that have no reportable spills or permit denials or shut-ins during our sample period.

† McFadden Adjusted R^2 for negative binomial and Probit models

Table 7. Results of Models with KLD Ratings for Total Environmental Strengths & Total Environmental Concerns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Log pounds emissions		Log regulatory penalty value		Number of regulatory penalties		Any major spills		Any permit denials or shut-ins	
	OLS		OLS		Negative binomial		Probit		Probit	
Total environmental strengths	0.068	0.008	0.039	0.038	0.025	0.024	0.016	0.014	0.009	0.005
	[0.129]	[0.062]	[0.071]	[0.060]	[0.044]	[0.037]	[0.010]*	[0.009]	[0.007]	[0.006]
Total environmental concerns	1.047	0.286	0.381	0.290	0.167	0.092	0.026	0.021	-0.002	-0.001
	[0.101]***	[0.051]***	[0.067]***	[0.053]***	[0.025]***	[0.021]***	[0.007]***	[0.007]***	[0.006]	[0.004]
Log assets, 2 years ago (+1,000)	0.100	0.131	0.327	0.274	0.105	0.109	0.057	0.049	0.018	0.006
	[0.224]	[0.111]	[0.125]***	[0.108]**	[0.070]	[0.054]**	[0.014]***	[0.013]***	[0.007]**	[0.006]
Log sales, 2 years ago (+1,000)	0.481	0.093	-0.098	-0.110	0.037	-0.013	-0.024	-0.020	-0.014	-0.003
	[0.235]**	[0.117]	[0.134]	[0.118]	[0.074]	[0.059]	[0.014]*	[0.014]	[0.008]*	[0.007]
Lagged dependent variable		Included		Included		Included		Included		Included
Year dummies	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
Industry dummies (3-digit NAICS Codes)	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
Observations (firm-years)	3742	3304	3137	2873	3115	2862	2762	2587	935	842
Firms	744	644	534	451	524	446	431	382	196	165
Adjusted R^2 †	0.499	0.738	0.322	0.378	0.133	0.169	0.310	0.332	0.408	0.500

All independent variables and controls are 2-year lags. For Negative Binomial and Probit models, this table reports marginal effects, not coefficients. Brackets contain standard errors clustered by firm. * significant at 10%; ** significant at 5%; *** significant at 1%. The samples in the Probit models are smaller because these models drop observations from industries that have no reportable spills or permit denials or shut-ins during our sample period.

† McFadden Adjusted R^2 for negative binomial and Probit models

Table 8. Results of Models with All Disaggregated KLD Environmental Ratings

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Log pounds emissions		Log regulatory penalty value		Number of regulatory penalties		Any major spills		Any permit denials or shut-ins	
Functional form	OLS		OLS		Negative binomial		Probit		Probit	
KLD Strengths (each 0 or 1)										
Beneficial products & services	0.922	0.363	0.290	0.228	0.292	0.242	0.110	0.095	0.012	0.019
	[0.293]***	[0.130]***	[0.120]**	[0.104]**	[0.137]**	[0.100]**	[0.039]***	[0.034]***	[0.025]	[0.025]
Pollution prevention	-0.208	-0.197	0.005	0.014	0.024	0.033	-0.022	-0.019	0.011	-0.003
	[0.244]	[0.129]	[0.144]	[0.123]	[0.085]	[0.063]	[0.014]	[0.015]	[0.016]	[0.005]
Recycling	0.512	0.091	0.394	0.296	0.023	0.055	0.001	0.001	-0.008	-0.009
	[0.361]	[0.179]	[0.262]	[0.218]	[0.112]	[0.089]	[0.025]	[0.024]	[0.008]	[0.004]**
Alternative fuels	-0.850	-0.346	-0.273	-0.129	-0.063	-0.022	0.035	0.057	0.060	0.043
	[0.505]*	[0.243]	[0.200]	[0.168]	[0.093]	[0.078]	[0.023]	[0.026]**	[0.036]*	[0.027]
Communications	0.165	0.186	-0.346	-0.315	-0.064	-0.081	0.004	-0.016		
	[0.256]	[0.199]	[0.149]**	[0.134]**	[0.101]	[0.089]	[0.025]	[0.020]		
Property, plant and equipment	-0.663	-0.310	-0.643	0.178	-0.239	-0.177	-0.006	-0.060		
	[0.760]	[0.349]	[0.643]	[1.139]	[0.170]	[0.192]	[0.054]	[0.007]***		
Other strength	-0.216	-0.274	-0.001	-0.009	-0.048	-0.050	-0.008	-0.011		
	[0.228]	[0.155]*	[0.131]	[0.127]	[0.080]	[0.076]	[0.021]	[0.021]		
KLD Concerns (each 0 or 1)										
Hazardous Waste	1.019	0.437	0.290	0.211	0.112	0.033	0.039	0.040	0.024	0.019
	[0.237]***	[0.130]***	[0.157]*	[0.129]	[0.079]	[0.050]	[0.019]**	[0.017]**	[0.015]	[0.012]
Regulatory problems	0.521	0.236	0.405	0.287	0.270	0.198	0.011	0.001	-0.020	-0.016
	[0.212]**	[0.108]**	[0.132]***	[0.117]**	[0.076]***	[0.054]***	[0.016]	[0.014]	[0.008]**	[0.006]***
Ozone depleting chemicals	1.043	0.891	0.661	0.397	-0.003	-0.147	0.214	0.201		
	[0.374]***	[0.247]***	[0.394]*	[0.295]	[0.169]	[0.109]	[0.115]*	[0.099]**		
Substantial emissions	1.646	0.190	0.584	0.496	0.121	0.076	0.041	0.038	0.018	0.012
	[0.249]***	[0.144]	[0.160]***	[0.146]***	[0.066]*	[0.050]	[0.019]**	[0.018]**	[0.018]	[0.013]
Agricultural chemicals	1.149	0.309	1.096	0.817	0.800	0.319	0.096	0.070	-0.011	-0.008
	[0.434]***	[0.229]	[0.290]***	[0.238]***	[0.38]**	[0.145]**	[0.095]	[0.072]	[0.005]**	[0.003]***
Climate change	2.037	0.349	-0.677	-0.472	0.112	0.097	-0.031	-0.030	-0.004	-0.002
	[0.439]***	[0.228]	[0.185]***	[0.164]***	[0.103]	[0.099]	[0.014]**	[0.015]**	[0.007]	[0.005]
Other concern	0.163	-0.032	0.127	0.167	0.100	0.059	0.002	-0.003	0.092	0.016
	[0.516]	[0.232]	[0.213]	[0.196]	[0.096]	[0.067]	[0.025]	[0.025]	[0.093]	[0.026]
Log assets	0.146	0.161	0.307	0.253	0.158	0.116	0.054	0.044	0.005	-0.001
	[0.227]	[0.111]	[0.118]***	[0.104]**	[0.075]**	[0.054]**	[0.013]***	[0.013]***	[0.006]	[0.005]
Log sales	0.458	0.075	-0.075	-0.075	-0.012	-0.009	-0.019	-0.013	-0.005	0.002
	[0.241]*	[0.118]	[0.127]	[0.113]	[0.080]	[0.058]	[0.013]	[0.013]	[0.007]	[0.006]
Lagged dependent variable		Included		Included		Included		Included		Included
Year dummies	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
Industry dummies (3-digit NAICS Codes)	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
Observations (firm-years)	3,743	3,305	3,138	2,874	2,855	2,855	2,763	2,588	905	813
Firms	744	644	534	451	441	441	431	382	192	162
Adjusted R^2 †	0.517	0.740	0.340	0.387	0.139	0.170	0.318	0.339	0.458	0.522
Wald test: KLD strength subscores jointly = 0? ††	2.35**	2.36**	1.92*	1.56	9.82	11.26	22.15***	39.65***	14.2**	18.37***
Sum of KLD strength subscore coefficients†††	-0.34	-0.49	-0.57	0.26	6.89	7.01	0.11	0.05	0.08	0.05
Wald test: Sum = 0? ††	0.09	0.88	0.56	0.05	0.27	0.06	0.72	1.35	1.32	0.04
Wald test: KLD concern subscores jointly = 0? ††	30.18***	9.86***	12.67***	9.37***	94.72***	34.98***	43.87***	44.22***	26.51***	16.59**
Sum of KLD concern subscore coefficients†††	7.58	2.38	2.49	1.90	9.72	8.18	0.37	0.32	0.10	0.02
Wald test: Sum = 0? ††	94.56***	34.14***	22.49***	20.29***	47.8***	12.47***	14.75***	11.94***	4.91***	3.66***

For Negative Binomial and Probit models, this table reports marginal effects, not coefficients. Brackets contain standard errors clustered by firm. * significant at 10%; ** significant at 5%; *** significant at 1%. All independent variables and controls are 2-year lags. The samples in the Probit models are smaller because these models drop observations from industries that have no reportable spills or permit denials or shut-ins during our sample period.

† McFadden Adjusted R^2 for negative binomial and Probit models †† Wald test results: F-value for OLS models, chi-squared value for negative binomial and Probit models.

††† Sum of coefficients for OLS models; sum of marginal effects for Negative Binomial and Probit models.

Table 9. Results of Models with Select Disaggregated KLD Environmental Ratings

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Log pounds emissions		Log regulatory penalty value		Number of regulatory penalties		Any major spills		Any permit denials or shut-ins	
Functional form	OLS		OLS		Negative binomial		Probit		Probit	
KLD Strengths (each 0 or 1)										
Pollution prevention	-0.215	-0.165	-0.009	-0.002	0.019	0.029	-0.021	-0.02	0.017	-0.001
	[0.239]	[0.124]	[0.139]	[0.116]	[0.085]	[0.062]	[0.014]	[0.015]	[0.019]	[0.008]
Property, plant and equipment	-0.525	-0.259	-0.689	0.121	-0.237	-0.172	-0.008	-0.063		
	[0.741]	[0.336]	[0.667]	[1.152]	[0.0175]	[0.197]	[0.055]	[0.007]***		
KLD Concerns (each 0 or 1)										
Hazardous waste	0.929	0.414	0.364	0.26	0.151	0.037	0.05	0.049	0.02	0.017
	[0.246]***	[0.127]***	[0.159]**	[0.126]**	[0.093]	[0.053]	[0.022]**	[0.019]**	[0.016]	[0.013]
Regulatory problems	0.692	0.277	0.455	0.318	0.309	0.210	0.014	0.002	-0.027	-0.024
	[0.215]***	[0.108]**	[0.137]***	[0.117]***	[0.084]***	[0.057]***	[0.017]	[0.015]	[0.011]**	[0.009]**
Substantial emissions	1.935	0.251	0.632	0.521	0.149	0.080	0.045	0.037	0.009	0.011
	[0.258]***	[0.143]*	[0.170]***	[0.150]***	[0.075]**	[0.051]	[0.020]**	[0.018]**	[0.015]	[0.014]
Log assets	0.037	0.129	0.297	0.245	0.146	0.109	0.057	0.048	0.015	0.005
	[0.231]	[0.109]	[0.123]**	[0.106]**	[0.079]*	[0.057]*	[0.014]***	[0.013]***	[0.006]**	[0.005]
Log sales	0.603	0.106	-0.067	-0.078	0.006	-0.005	-0.022	-0.018	-0.013	-0.003
	[0.244]**	[0.116]	[0.134]	[0.117]	[0.085]	[0.061]	[0.014]	[0.014]	[0.007]*	[0.006]
Lagged dependent variable		Included		Included				Included		
Year dummies	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
Industry dummies (3-digit NAICS Codes)	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
Firm year observations	3743	3305	3138	2874	2858	2858	2763	2588	936	843
Firms	744	644	534	451	442	442	431	382	196	165
Adjusted R^2 †	0.49	0.74	0.32	0.38	0.14	0.17	0.31	0.33	0.43	0.51
Wald test: Strengths jointly zero ? ††	0.65	1.20	0.54	0.01	1.14	0.75	1.77	23.47***	n/a	n/a
Wald test: Sum of strengths equal zero? ††	0.90	1.41	1.08	0.01	0.87	0.37	0.28	21.95***	n/a	n/a

For Negative Binomial and Probit models, this table reports marginal effects, not coefficients. Brackets contain standard errors clustered by firm. * significant at 10%; ** significant at 5%; *** significant at 1%. All independent variables and controls are 2-year lags. The Probit models drop observations from industries that have no reportable spills or permit denials or shut-ins during our sample period. *Property, Plant and Equipment* is dropped from Models 9 and 10 because it perfectly predicts no permit denials/shut-ins.

† McFadden's adjusted R^2 for negative binomial and Probit models.

†† Wald test results: F-value for OLS models, chi-squared value for negative binomial and Probit models.

Table 10. Comparison of Model Fit Statistics

PANEL A

Emissions (OLS models)

			Difference Significance of difference
	Baseline 0.413	KLD 0.499	0.087 482.9***
	Lags 0.832	Full 0.832	0.001 24.5**
Difference	0.419	0.333	
Significance of difference	3677.7***	3219.4***	

Penalty value (OLS models)

			Difference Significance of difference
	Baseline 0.325	KLD 0.368	0.043 194.8***
	Lags 0.453	Full 0.459	0.006 42.8***
Difference	0.128	0.091	
Significance of difference	582.3***	430.3***	

PANEL B

Number of regulatory penalties (Negative binomial models)

			Difference Significance of difference
	Baseline 0.132	KLD 0.144	0.012 n.s.
	Lags 0.191	Full 0.192	0.001 n.s.
Difference	0.062	0.049	
Significance of difference	♦♦♦	♦♦♦	

Any major spills (Probit models)

			Difference Significance of difference
	Baseline 0.290	KLD 0.320	0.029 n.s.
	Lags 0.370	Full 0.375	0.005 n.s.
Difference	0.079	0.055	
Significance of difference	♦♦♦	♦♦♦	

Any permit denials or shut-ins (Probit models)

			Difference Significance of difference
	Baseline 0.462	KLD 0.480	0.018 n.s.
	Lags 0. 532	Full 0.524	-0.008 n.s.
Difference	0.070	0.044	
Significance of difference	n.s.	n.s.	

Table cells present a fit statistics for nested regression models. The baseline model includes log assets, log sales, and industry dummies (3-digit NAICS codes). The KLD model adds the KLD subscores to the baseline model. The Lags model adds 1-, 2-, and 3-year lags of the dependent variable and 2-year lags of the other outcome variables to the baseline model. The full model adds both sets of variables to the baseline model. In Panel A, the cells contain Adjusted R^2 values and the surrounding figures display the difference in these values and the LR chi-squared test statistic indicating whether these differences are statistically significant with * p<0.10; ** p<0.05 *** p<0.01. In Panel B, the cells contain McFadden's Adjusted R^2 values for these negative binomial and Probit models. Figures outside these cells display the difference in these values and the test statistic that the differences in the Bayesian information criteria (BIC) between these models are statistically significant. We employ comparison guidelines proposed by Raftery (1995), where larger absolute differences of BIC provide increasing evidence favoring the more full complete model (the extent of evidence provided by differences of 0-2 is "weak", 2-6 "positive", 6-10 "strong", and >10 "very strong"). We denote the latter three categories as ♦ "positive", ♦♦ "strong", and ♦♦♦ "very strong". n.s. denotes that the test lent support for the more limited model.

Appendix A: Modeling the Social Rater's Problem

This appendix presents an illustrative model of the social rater's problem in summarizing multiple metrics into an optimal index of social or environmental performance to provide transparency to stakeholders. The social rater must construct its rating $K_{f,0}$ for firm f in year 0 from two pieces of data: its measure of the firm's management quality measured in year 0 ($m_{f,0}$) and information available in year 0 on this firm's historical environmental performance ($Y_{f,0}$). Management quality refers to its ability to make investments that improve environmental performance and reduce risks of environmental problems. We normalize the measure of management quality m to have the same units as environmental performance Y and mean of zero. We assume the rater uses a linear rule, meaning that the social score depends on these two factors with a weight α chosen optimally by the social rater:

$$K_{f,0} = \alpha\, m_{f,0} + (1-\alpha)\, Y_{f,0} \qquad \textbf{(A1)}$$

Future environmental performance is determined by both management quality and by whatever determined historical environmental performance; hence, both management quality and historical environmental performance are useful in predicting future emissions:

$$Y_{f,1} = c_1\, m_{f,0} + c_2\, Y_{f,0} + v_f \qquad \textbf{(A2)}$$

Equation A2 is not necessarily causal; it is just the reduced form that is the best linear predictor of future environmental performance.

As described in the text, we assume customers of the social rater desire transparency about firms' historical environmental performance $Y_{f,0}$ and future environmental performance $Y_{f,1}$. Thus, we assume the objective for the social rater is to minimize the average across firms of the weighted average squared deviation between the rating and environmental performance in each period:

$$\text{Argmin}(\alpha):\ \Sigma_f\, [\beta\, (K_{f,0} - Y_{f,0})^2 + (1-\beta)\, E_0(K_{f,0} - Y_{f,1})^2] \qquad \textbf{(A3)}$$

The parameter β is the social rater's summary of what it perceives customers want in terms of summarizing the past (β near 1) versus predicting the future (β near zero). The expectation E_0 is taken at time 0.

This illustrative model yields our first result:

> **Theorem 1:** The optimal weight α for the social rater to apply to its measure of management quality $m_{f,0}$ is:
>
> $$\alpha^* = (1-\beta) * (1-c_2)$$

Theorem 1 implies that the social rater pays more attention to management quality when the social rater perceives that its customers care more about the future (that is, β is low) and when emissions Y are not strongly predicted by history, even conditioning on the rater's measure of management quality (c_2 is low).

Sketch of the Proof: Theorem 1 follows from the social rater's first-order condition for optimizing its goal (equation 1); that is, setting the derivative of the social rater's objective (1) with respect to its choice variable α equal to zero.

This result has two corollaries. First, the boundary condition $\beta = 0$ means the social rater is only interested in summing up the past. That boundary condition implies that the rater sets its social rating equal to the past, $K_{f,1} = Y_{f,0}$.

Corollary 1: When the social rater cares only about the past ($\beta = 0$), the R^2 of the regression of the optimal rating on the future should equal the R^2 of the regression using historical data.[36]

When the social rater cares about predicting the future as well as summarizing the past (as is suggested by our transparency analysis in Section 2 of the paper), it can sacrifice some fit in the first term of equation A3 $(K_{f,1} - y_{f,0})^2$ to increase fit in the second term $E_0(K_{f,1} - y_{f,1})^2$. This possibility leads to:

Corollary 2: When the social rater cares about the future as well as the past, the R^2 of the regression of the optimal rating on the future should be higher than the R^2 of using only historical data to predict the future.

ENDNOTES

[1] For example, consider the detailed "Global Profiles" offered by the Sustainable Investment Research International (SiRi) Company, a consortium of ten socially responsible investment research organizations based in Europe, North America, and Australia, including Kinder, Lydenberg, Domini Research & Analytics (KLD). SiRi is "the world's largest independent provider of SRI research and consulting services for institutional investors and financial professionals." Each SiRi profile details historical environmental performance (e.g., energy and water consumption, wastes and emissions, compliance penalties), recently implemented environmental management activities (e.g., the percentage of a firm's plants certified to the ISO 14001 Environmental Management System Standard), and ongoing environmental activities (e.g., frequency of environmental audits) and environmental objectives and plans. *See* www.siricompany.com/services.shtml.

[2] Written statement by Raymond W. McDaniel, President of Moody's Investors Service http://www.sec.gov/news/extra/credrate/moodys.htm, accessed April 25, 2007.

[3] Caroline Wilbert, "Social Responsibility of Coke Questioned," *Atlanta Journal Constitution*, July 19, 2006.

[4] *Campaign to Stop Killer Coke,* "News Release: Coca-Cola Suffers Big Blow in Investment Community," July 18, 2006, http://killercoke.org/nr060718.htm, accessed September 19, 2006.

[5] Paul Hawken, "Socially Responsible Investing," The Natural Capital Institute, October 2004.

[6] This portion of our assessment is akin to research that investigates whether credit ratings such as Moody's and Standard & Poor's (S&P) provide transparency to investors by accurately predicting defaults on corporate debt (Güttler, 2005; Krämer and Güttler, 2006) and sovereign debt and currency crises (Mora, 2006; Sy, 2004). For a review of this literature, see Cantor (2004).

[7] Jon Entine, Academy of Management's Organizations and the Natural Environment Division listserv (ONE-L), September 14, 2006.

[8] This figure includes tobacco, alcohol, gambling, defense/weapons, community relations, environment, labor relations, products/services, equal employment, faith-based, pornography, human rights, animal testing, abortion, medical ethics, youth concerns, anti-family entertainment and lifestyle, and excessive executive compensation.

Social Investment Forum, 2006. *2005 Report on Socially Responsible Investing Trends in the United States* [http://www.socialinvest.org/areas/research/trends/SRI_Trends_Report_2005.pdf, accessed March 7, 2007]

[9] Note that we do not endorse any of the beliefs that underlie these motivations to invest in social funds. For example, although consequentialist investors believe they can shift the cost of capital in ways that reward good behavior, we are aware of no empirical evidence that supports this point.

[10] We are not asserting that such investors' beliefs are correct, merely that some investors exhibit them. Many of the more than 100 empirical studies that have investigated the relationship between CSR and financial performance (see Margolis and Walsh, 2003) have found a positive association (e.g., Orlitzky, Schmidt, and Rynes, 2003; Russo and Fouts, 1997). Careful studies of how CSR influences financial performance have acknowledged the important potential problem of reverse causality (e.g., Waddock and Graves, 1997). In addition, variables that might affect both CSR and financial performance should not be omitted.

[11] Ethical Investment Research Services, "A Brief History of SRI/ Ethical Investment" (http://www.eiris.org/pages/top_menu/key_facts_and_figures/history_of_ethical_investment.htm, accessed May 3, 2007).

[12] For another example of expressive motives for investment, see Trillium Asset Management at http://www.trilliuminvest.com/pages/sri/sri_home.asp, accessed April 17, 2007.

[13] KLD initially rated approximately 650 companies listed on the S&P 500 and the Domini Social 400 Index, but expanded its scope to all members of the Russell 1000, which added approximately 500 rated companies (Email communication on July 5, 2007 with Jay Carberry, KLD Research & Analytics, Inc.).

[14] Because of annual changes in the memberships of the stock indices for which KLD provides ratings, some firms entered and some exited our sample during the study period.

[15] KLD Research & Analytics, Inc., www.kld.com, accessed September 11, 2006.

[16] KLD coded three of these variables (*Communications*; *Property, plant and equipment*; and *Climate change*) for only a subset of the years we analyze. In our models, we recoded the missing values to zero and included dummy variables to denote recoded observations.

[17] Although it is common practice to aggregate them, "strengths" and "concerns" for a single CSR domain can represent distinct constructs (Mattingly and Berman, 2006). Moreover, that *total KLD environmental strengths* and *total KLD environmental concerns* are *positively* correlated (Mattingly and Berman, 2006) suggests that aggregation might cloak important differences: a firm with five KLD strengths and five KLD concerns is surely different from a firm with only one of each, a distinction lost in the summing of strengths and concerns.

[18] This includes regulations pursuant to the following statutes: the Atomic Energy Act; Clean Air Act; Clean Water Act; Endangered Species Act; Federal Insecticide, Fungicide, and Rodenticide Act; Mine Safety and Health Act; Resource Conservation and Recovery Act; Safe Drinking Water Act; and Toxic Substances Control Act.

[19] Our results were robust to adding 1 or 10 (instead of 1,000) to these variables.

[20] Spill data from the Corporate Environmental Profiles Directory database includes only oil spills in excess of 10,000 gallons and chemical spills in excess of 10,000 pounds.

[21] IRRC counts every instance in which a permit was denied at a given company facility regardless of whether the facility was ultimately granted a permit. IRRC (2002) notes that "Permit denials may result in financial losses: a company that operates a commercial hazardous waste treatment, storage or disposal facility cannot legally accept new business at the site until it receives a permit; a company that manages its own waste may be forced to reduce production levels temporarily or ship hazardous waste off-site at additional cost."

[22] IRRC (2002) notes that shut downs represent "the best available indicator of the safety of the operation of offshore oil platforms."

[23] Environmental regulators' delays in releasing data to the public range from a few weeks to as much as two years. For example, spill data are made publicly available by the Emergency Response Notification System (ERNS) within weeks of the occurrence of a spill. Most EPA compliance data are made available to the public within three months via databases such as Enforcement & Compliance History Online (ECHO) or procured via Freedom of Information Act requests. In contrast, Toxic Release Inventory emissions data are made available to the public two years after the emissions occur (e.g., 2004 TRI data was released in April 2006). To reflect these varying delays, we predict KLD ratings based on compliance, spills, and permit denials lagged one and two years (e.g., 2000 and 2001 compliance data to predict 2002 ratings) and TRI emissions lagged three and four years (e.g.,

1998 and 1999 TRI data to predict 2002 ratings, the TRI data having been made publicly available in 2000 and 2001).

[24] We use a Poisson model (rather than a negative binomial model) because *total environmental concerns* exhibited only limited overdispersion (mean=0.64, variance=1.08) and a likelihood-ratio test provided no evidence that alpha significantly differed from zero (chi-bar^2=3.2e-06, p=0.50). The results were unchanged when we used a negative binomial model, which drops the Poisson distribution's restrictive assumption that the mean equals the standard deviation.

[25] As noted earlier, TRI emissions are made publicly available two years after the emissions occur. Thus, we predict TRI emissions that occur in a given year based on KLD ratings issued that same year (e.g., 2000 KLD ratings to predict 2000 TRI emissions, which are not made publicly available until 2002).

[26] Including a lagged dependent variable is a common way to control for such autocorrelation (Keele and Kelly 2006)

[27] We approximate these effects with marginal effects based on derivatives, with all other variables set at their mean.

[28] Because KLD might consider additional environmental harms that are not in our data, we are unable to look for false positives (that is, firms that did not deserve the *concerns* rated by KLD).

[29] This marginal effect varies substantially depending on firms' penalty value. For example, consider two firms whose penalty value is two standard deviations above the sample mean. Our results imply that the one with one fewer (more) KLD environmental strength (concern) will experience a 29% ($39,500) increase in penalties.

[30] This marginal effect varies substantially depending on firms' penalty value. Among firms with penalties two standard deviations above the sample mean, one additional KLD environmental concern is associated with a subsequent 47% ($63,000) increase in penalties.

[31] These ratios were calculated as net income per total assets, net income per total common equity, and net income per total net sales, respectively. We obtained these data from Compustat. We recoded values of these ratios beyond the 1st or 99th percentiles to these threshold limits to reduce the impact of outliers.

[32] We log each ratio after adding 1,000.

[33] We thank an anonymous reviewer for this suggestion.

[34] See, for example, the more than 100 studies summarized in Dawes, Faust, and Meehl (1989).

[35] In 2005, after the period analyzed in this article, this issue was incorporated into the Corporate Governance Transparency rating.

[36] R^2 and adjusted R^2 are equivalent in this case because there is one rating and one dimension of measurement. As long as the number of ratings equals the number of dimensions, this condition will hold.