Evaluating companies’ social and environmental performance: Current practice and some recommendations

Ted Martin Hedesström and Anders Biel

Hedesström, T. M., & Biel, A. Evaluating companies’ social and environmental performance: Current practice and some recommendations. Göteborg Psychological Reports, 2008, 38, No. 1. As a tool for sustainable and socially responsible investing (SRI), an increasing number of specialist analyst agencies rate companies’ social and environmental performance and sell that information on to investors. A comparison of the most commonly engaged SRI rating agencies among Swedish institutional investors shows that they differ with respect to what criteria they use when assessing companies. While this may reflect agencies’ efforts to retain a competitive edge, it also indicates that knowledge about the relative relevance of particular performance indicators on financial, environmental, and societal risk is not yet well-advanced. A higher degree of transparency into the practices of SRI analyst agencies, as well as increased cooperation between SRI analysts and the scientific community, is warranted in order to assess and improve the reliability and validity of the social and environmental performance measures currently in use.

Key words: Sustainable investment, company assessment, validity, reliability

This research was supported by grants from MISTRA to the programme Behavioural Impediments to Sustainable Investment. We are grateful to Sasja Beslik, head of Responsible Investments at Banco; Magnus Furugård, CEO of GES Investment Services; and Anna Nilsson, head of SRI Analysis at Robur, whom we interviewed for the purpose of this article. Any opinions, conclusions or recommendations expressed are those of the authors.
How to make corporations adopt environmentally and socially sustainable practices has been a topic of academic research for several decades. The emphasis of much of the empirical studies within this area has been on seeking to establish that there exists a positive relationship between companies’ environmental/social conduct (often simply labelled corporate social performance, CSP) and their financial performance (CFP). The argument is that if such a link could be determined, this would make a persuasive case for companies to take environmental and social issues more seriously (Rennings, Schröder, & Ziegler, 2006; Wood & Jones, 1995).

Recently the role of investors in improving CSP has received increased attention. How can investors exert pressure on corporations to better their environmental and social conduct? The rationale behind sustainable and socially responsible investing (SRI) has hitherto largely relied on the “cost of capital” argument (Haigh & Hazelton, 2004; Statman, 2000), which implies that by excluding unethical companies from their investment portfolio while including ethical companies, SRI funds are able to reduce the cost of capital for “good” relative to “bad” companies, thereby instigating corporate change. However, since SRI so far largely has been confined to specialist retail niche funds, representing only a small share of the total fund market, the effects of these funds’ investments and disinvestments on corporations’ operations have been negligible (Haigh & Hazelton, 2004). Furthermore, while SRI funds often avoid investing in some particular cherry-picked “sin” industries, such as alcohol, tobacco, and pornography (Bauer, Koedijk, & Otten, 2005), other aspects of companies’ social and environmental conduct may not be taken into account.

In order for sustainable investments to have a substantial impact, many have argued that not only SRI retail funds, but also large mainstream institutional investors such as pension funds and insurance companies, will need to integrate environmental and social performance criteria into their analyses of companies (e.g., Hawley & Williams, 2000; Sparkes & Cowton, 2004). SRI should furthermore move from being a screening-avoidance paradigm to becoming a comprehensive paradigm that also seeks to affect corporate behaviour by, for example, entering into dialogue with companies that fall below certain CSP standards, and/or encouraging good practices by use of best-in-class investment approaches where only companies with the best CSP in each given industry are available for inclusion in the fund’s investment portfolio (Mackenzie, 2006). In this new vintage of SRI, CSP analysis should not only focus on potential risks associated with below-par social and environmental practice, but also identify business opportunities associated with good practice (Figge & Hahn, 2006; Sullivan & Mackenzie, 2006).

While some argue that mainstreaming of sustainable investing is already under way (Collier, 2004; McCann, Solomon, & Solomon, 2003), others claim there is still a long way until institutional investors integrate CSP analysis into their core investment decision processes (Dillenburg, Green, & Erekson, 2003; Lydenberg, 2005; Mackenzie & Sullivan, 2006). A major factor impeding such a development is the lack of standardization with regards to how companies report extra-financial information, and the difficulty for analysts to obtain such information and make use of it (Beloe, Scherer, & Knoepfel, 2004). The quality of information available to investors and analysts about companies’ social and environmental performance is generally poor (Merme & Zedek, 2004; Sethi, 2005), and extra-financial information presented in company reports and other official documentation is often of essentially qualitative nature. This poses a problem, since non-SRI financial analysts are reluctant to use information that is not quantified (Guyatt, 2006; Hassel, 2006). Dillenburg et al. (2003) conclude that the reason why CSP assessment is not yet a mainstream component
of investment analysis is that metrics that capture social and environmental management processes are still underdeveloped. They propose that once rating schemes measuring CSP are well-developed and coordinated, such schemes will have “a tremendous impact in influencing corporate behaviour … (since) what gets measured gets managed” (Dillenburg et al., 2003, p. 169). The problem, then, is how to “translate” firms’ social and environmental performance into standardised and comparable quantitative measurements. The next section describes what efforts are being made to this end, as well as criticism directed against such attempts.

Measuring corporate social and environmental performance

Wood (1991) defines CSP as “a business organization’s configuration of principles of social responsibility, processes of social responsiveness, and policies, programs, and observable outcomes as they relate to the firm’s societal relationships” (p. 693). In this oft-quoted and widely accepted definition, CSP thus refers both to a corporation’s processes – its organizational structure, operations, and management practices – and the consequences of its activities for stakeholders (Husted, 2000; Preston, 1990).

The last three decades have seen the emergence of a vast body of empirical studies investigating the afore-mentioned elusive link between companies’ social/environmental and financial performance. Because of the problems in handling the many and multifaceted variables that may affect this relationship (Carroll, 2000), most of these studies focus on a single dimension, such as, for example, air pollution, philanthropy, illegal activity, employee relations, or product recall, in their operationalisation of CSP (Cox, Brammer, & Millington, 2004; Waddock & Graves, 1997). Disappointingly, the collective evidence from this research could best be described as inconclusive (Griffin & Mahon, 1997; Wood & Jones, 1995). Rowley and Berman (2003) provide an explanation for this by convincingly argue that attention to any one stakeholder group, or issue, neglects other groups/dimensions and cannot independently serve as a proxy for CSP. Single dimension research designs therefore lack the reliability and validity necessary for making generalisations regarding the CSP-CFP link.

One way to evade the problems associated with single dimension CSP-CFP studies is to use a composite measure, where a firm receives a score based on the sum or average tallies across several social and environmental dimensions. Such ratings are to an increasing extent being made available within the investment community, thanks to the recent and manifold emergence of SRI rating agencies that evaluate companies’ social and environmental performance and sell that information on to institutional investors (European Commission, 2000). The aggregated CSP measures created by these specialised SRI analyst teams are being used by many academic researchers investigating the relationship between firms’ social/environmental and financial performance (e.g., Sauer, 1997; Waddock & Graves, 1997). However, the reliability and validity of these ratings is frequently questioned on theoretical and methodological grounds.

One objection relates to the apparent lack of theory behind the selection of criteria on which the CSP measures are based. According to Entine (2003), there are no agreed-upon standards or theoretical rationale underlying the way SRI rating agencies aggregate and compare multiple dimensions across or within industries. Sethi (2005) similarly maintains that the selection of criteria on which the CSP measurements are based are “almost invariably subjective,” and that these measurements, as currently practiced, are “at a rudimentary state and do not compare well in terms of conceptual rigor and methodological sophistication of more traditional financial criteria of corporate performance” (p. 108). Rowley and Berman
No. 1:38, 4

(2000) argue that the rating agencies’ aggregation of multiple social and environmental data is not supported by an underlying understanding of CSP, and illustrate the problems involved in aggregating various social and environmental attributes into one single composite CSP measure: “What does a score of 6 mean? How do you compare a firm that receives a satisfactory rating on all dimensions with a firm that receives poor ratings for half the dimensions and excellent ratings for the other categories? … (A) firm that treats all of its stakeholders “reasonably” well may receive a similar rating to a firm that is well above average in its employee policies but is well below average on pollution abatement” (Rowley & Berman, 2000, p. 403).

Furthermore, since different types of social and environmental activities may have different implications for a firm’s financial performance (Griffin & Mahon, 1997; Hillman & Keim, 2001), it has been questioned whether composite CSP measures can be used for drawing conclusions regarding the CSP-CFP link. Empirical evidence indicates that various CSP dimensions are not strongly correlated with regard to their relationship with financial performance (Berman, Wicks, Kotha, & Jones, 1999; Cox et al., 2004; Hart, 1995; Johnson & Greening, 1999; Wood & Jones, 1995). Although rating agencies often create separate indices for social, environmental, and economic (governance) dimensions of CSP, this only partly alleviates the problem since even within each CSP dimension, different attributes may have different financial consequences. Lankoski (2006) argues, for example, that particular environmental issues vary when it comes to their link with economic performance, as they involve different options for solutions, give raise to different regulatory efforts, and provoke different reactions in customers and other stakeholders. The precariousness of weighting different social and environmental aspects against each other deter some rating agencies from applying differentiated weightings at all in their composite CSP measure; instead they assign equal weighting to each attribute (Waddock & Graves, 1997).

Another strand of criticism of the CSP ratings targets the methods by which data about companies is collected. Entine (2003) sketches the following unflattering picture of the activities of some rating agencies, among those the US SRI analyst firm Kinder, Lydenberg, & Domini (KLD), provider of the Domini 400 Social Index: “(They) review thousands of companies with skimpy resources. Research relies upon often unreliable (and) anecdotal … data. Overworked and undertrained junior staffers draw on government data banks, journalistic sources, and information supplied by companies, collecting whatever information they select as relevant. The task is daunting and highly subjective. These data are then given a patina of objectivity by being turned into hierarchical numbers” (Entine, 2003, p. 355). A report from SustainAbility and MISTRA, investigating the screening methodology of 11 European and four North American SRI rating agencies, confirms that the analysts’ qualifications are not always up to standard. The proportion of SRI analysts with financial or business experience is, the report states, often low, and the analyst teams therefore “typically lack the skills needed to address financial and strategic considerations” (Beloe et al., 2004, p. 2).

Some rating agencies draw predominantly on company reports and other official documentation as the basis for their CSP assessment. This method is likely to carry an inherent bias against smaller companies. Large companies employ staff – or perhaps entire departments – who deal exclusively with writing up reports and supplying information via the company’s web page, while small companies, because of lack of resources, generally provide less information of this kind. When rating agencies go through companies’ official documentation in search for specific social and environmental information, lack of data on
any of the criteria of interest automatically renders the company a lower over-all CSP score. While this mechanism serves as a useful incentive for companies to include extra-financial information in their company reports, larger companies tend so far, as a result, to receive more favourable CSP ratings than smaller companies (M. Furugård, CEO of GES Investment Services, personal communication, June 2, 2006).

An alternative method of collecting social and environmental information is through questionnaires which companies are asked to fill out. This method is quite common: in a survey of 14 European and one US SRI analyst firms, 11 reported that they were – at least partly – relying on questionnaires as a basis for their ratings (Franck & Böcke, 2003). The use of questionnaires has a distinct advantage compared to other methods of collecting extra-financial information, as it allows for standardized reporting across all investigated firms. Since, as a rule, each rating agency constructs its own questionnaire, increasing demand is however being placed on targeted companies’ internal resources to answer a multitude of surveys of shifting quality, which threatens to result in “questionnaire fatigue” (Beloe et al., 2004). As representatives of one company (British American Tobacco) put it, “(s)ome of these questionnaires are neither relevant to the sector nor acknowledge that the data is in the public domain through published reports” (Edmondson & Payne, 2006, p. 251). The Green Paper “Promoting a European framework for CSR,” issued by the European Commission in 2001, echoes these concerns. As a consequence of the many SRI analyst agencies using a number of different tools and metrics, companies seem overloaded with excessive and divergent information requests. The report warns that a further expansion of SRI may, as a result, encounter a growing aversion and non-cooperation from companies. To remedy this, the Commission warrants “further standardisation, harmonisation and transparency in screening tools and metrics used by screening agencies” (European Commission, 2001, p. 22). In a response to this request, 15 European SRI analyst firms agreed by the end of 2001 to develop a Voluntary Quality Standard (VQS). While aiming at developing a “standard of SRI research and scope,” it is notable that the VQS initiative at the same time claims to be “trying not to standardise or harmonise criteria” (Franck & Böcke, 2003, p. 4). This appears to reflect an inherent resistance within the SRI research community against too far-reaching standardization attempts. As it is in the interest of each analyst firm to claim the superiority of its own rating scheme, it is not surprising that they often remain largely secretive about their exact methodology (Hoggett & Nahan, 2002).

The present investigation is aimed at illuminating to what extent various SRI analyst organisations engaged by Swedish institutional investors use common criteria when assessing companies’ social and environmental performance; what conventions, guidelines, or international agreements they claim to be the basis from which these criteria are developed; and what methodology they use in collecting and presenting data. With this comparison as a point of departure, we will discuss how analysts, index providers and independent organisations setting standards for how Swedish institutional investors assess companies’ social and environmental performance could address some of the issues raised in the literature.

As a first step in the investigation, we surveyed websites of Swedish intuitions that claim to invest in accordance with social, environmental, and ethical criteria, in order to find out what those criteria are and how extra-financial information about companies is collected and analysed. The survey confirmed that it is common practice to rely on external analysts for this purpose. There are exceptions to this general rule, but they mainly apply to the rating of companies listed in Sweden. The two most commonly engaged SRI analyst firms among
Swedish institutional investors are Global Ethical Standards (GES) and Innovest (formerly Coreratings) (Skillius, 2005). In addition, some institutions depend on the sustainability index providers Ethibel, FTSE4Good, and Sustainable Asset Management (SAM); and/or on the independent organisations Ethical Investment Research Services (EIRIS) and Global Reporting Initiative (GRI). In the following, we will compare the rating instruments used by these SRI analyst agencies. The objective is to identify the evaluation scheme that guides their analyses and review the extent to which they target the same issues. In making comparisons between agencies, we draw on information from official documents, web pages and, in some cases, interviews with analysts. We compare the specific environmental and social indicators the SRI analyst agencies use, the international standards and guidelines they refer to, and the methods by which they collect, evaluate and report extra-financial information about companies. We end this section by discussing the reliability and validity of the agencies’ methods and measures.

Corporate social and environmental assessment by SRI analyst agencies

Current practice

International conventions. As revealed by the table in Appendix, several analyst agencies allude to international conventions as the backbone of their analysis. Common among these are the United Nations Global Compact (UNGC; ten principles in the areas of human rights, labour, the environment and anti-corruption, derived from the Universal declaration of human rights, the International Labour Organization’s [ILO] declaration of fundamental principles and rights at work, the Rio declaration on environment and development, and the United Nations’ convention against corruption), the ILO core conventions, and the Organisation for Economic Co-operation and Development (OECD) guidelines for multinational enterprises. Some issues covered by these conventions are likely to be more easily operationalised than others. As an example, child labour is in many countries guided by national law, and a minimum age for admission to employment or work is specified. Laws against corruption are established in many countries, and it is also illegal in many countries for companies to engage in corrupt practices in another country. Other issues – not least, environmental issues – might be more difficult to come to grips with; for example principle seven in UNGC, the precautionary approach. One concept in this approach is preventive anticipation: “taking action if necessary before scientific proof is available on the grounds that a delay in action will cause damage to nature and society (United Nations Global Compact, 2006).” The less precise the concepts or issues to be converted into practical measures are, the more likely it is that analysts choose somewhat different specifications, which in turn may lead to differing evaluations of a company’s performance.

Assessment criteria. The specific social and environmental criteria that SRI analysts rely on (see Appendix) have somewhat different headings across the SRI analyst community. Some list social, environmental, governance, and ethical criteria, while others make use of fewer clusters. We have chosen to present the criteria under two main headings: ‘social’ and

---

1 Under the auspice of the United Nations Environment Programme (UNEP), GRI is an international, multi-stakeholder effort to create a common framework for voluntary reporting of the economic, environmental, and social impact of organisation-level activity. In contrast to the other SRI organisations investigated here, they do not compile ratings. Since some institutions claim to collect CSP data in accordance with GRI’s criteria, we will nevertheless include them in the comparison.

2 Comparisons will be made with anonymity preserved.
‘environmental.’ ‘Environmental’ is further divided into internal environmental policy, referring to the processes by which a firm manages environmental issues; and environmental performance, referring to outputs such as pollution and waste. ‘Social’ is similarly split up into internal social policy, referring to management practices and employee conditions; and social performance, referring to criteria associated with external stakeholder relations. Furthermore, to allow for comparisons across analyst agencies, we have arranged some particular criteria under more common labels, for example, management engagement.

As the Appendix shows, some criteria are commonly applied across analyst agencies, while others are less common. Many criteria concern internal social policy, strategy management, and human capital. Overall, the internal policy measures seem to be more homogeneously applied than the performance measures. Notwithstanding agreements, many cells in the table are left blank, indicating that the different agencies do not apply the same criteria.

Data collection. The various analyst agencies differ considerably with regards to the methods by which they collect information about companies. Most use several complementary methods. While all agencies examine official documents, such as company reports, three also send questionnaires for companies to fill out. Agencies that do not use questionnaires instead rely on interviews with company representatives and stakeholders, and on reports from non-governmental organisations (NGOs) and the media.

Data reporting. All of the analyst agencies in the sample claim to be sensitive to which industry or branch a company is engaged in. Typically, an industry could be classified on a three-level scale, where A denotes low risk, B medium and C high risk. Some agencies also make country-specific classifications into high, medium and low-risk countries.

All agencies report CPS ratings multidimensionally, that is, they rate each company separately on different dimensions, typically social, environmental, and economic (governance) dimensions. In addition to the industry (and/or country) risk rating, each company is thus rated for its particular risk level on each dimension. While some agencies apply differential weightings to the various attributes within a dimension, others weight each attribute equally. Three agencies also create a composite CPS measure, encompassing all dimensions.

Reliability or diagnosis

Analysts serve as experts for the institutional investors. Their role is to analyse alternative investments and describe possible outcomes for their clients. Weiss and Shanteau (2004) draw a parallel between the decision making of experts and medical judgements and distinguish between three levels of decisions. The first level is diagnosis, the second prognosis and the third level is treatment. In the present context analysis corresponds to the level of diagnosis, while description of possible outcomes tallies with prognosis. Based on this information institutions decide what to do; the treatment level. Several criteria have to be fulfilled in order for the diagnosis to be reliable. Experts should show a high degree of consensus, or interobserver agreement. As proposed by Einhorn (1974), experts must also be consistent. Given repeated assessments of the same stimulus, the diagnosis should be the same.

How could we know if analysts do agree? To the extent that they come up with the same rank order of companies with regard to social and environmental performance, this would indicate agreement. The same would be true for similar orders of industry or country
risk. Unfortunately, such agreement is hard to establish, given that (some) rating agencies do not reveal all their information. Agreement is also hard to establish since different rating agencies partly use different criteria. One way to circumvent these problems would be if analysts agreed to use a common set of criteria and make information available to independent bodies of valuers. Although new knowledge in the environmental and social domains may call for additional criteria to be used in the analysis, more important issues that fall back on international conventions should be of a more permanent character. This does not preclude that analysts keep information about certain idiosyncratic indicators to themselves.

Since analyses are made in a constantly changing world, consistency is not easily established. However, consistency is partly dependent on the same instrument being used in a consistent and compatible manner. This is more likely to be the case if questionnaires, or structured interviews with company representatives, are used. Less structured interviews will probably vary from time to another. Similarly, reports from external sources will also vary between sources and across time.

To sum up, reliability is served if a common set of criteria is applied in a systematic manner over time. However, reliability does not guarantee validity. Experts may all agree, and the same expert may be consistent from one time to another. Still, are the wrong criteria used while reliability is high, validity is doomed to be low. On the other hand, validity can only be high to the extent that reliability is also high.

Validity and prognosis

In the present context, validity concerns the relationship between the diagnosis and potential risks or outcomes. To what extent is the analyst’s advice about which companies to stay away from and which to invest in a valid one? A conceptual framework for the relationship between assessment of CSP and potential risks and outcomes is proposed in Figure 1. As can be seen, several such prognoses come into play. CSP analysis could be used in order to predict the financial risk of companies. In that case, environmental and social indicators are relevant to the extent that they predict financial risks. Presently, little is known about the validity of CSP rating instruments in predicting financial risks. However, investors may also – directly, or indirectly through fiduciaries – be interested in the social and environmental effects of company activities, regardless of their financial consequences. While CPS analysis may be used to this end, the question is whether the rating instruments currently in use properly capture these risks. Also here, little is known about the performance of rating instruments, but natural scientists could play an important role in establishing their validity.

The arrow between environmental and financial outcomes in Figure 1 indicates that even if environmental and social risk is not the investor’s prime interest, environmental and social outcomes may still affect financial outcomes. In particular, this concerns the aggregated effect of companies’ activities rather than the effect of the activities of a single company, the unit that analysts rate. Companies may deplete natural resources that constitute a necessary element of their enterprise, or stricter regulation may be imposed on companies due to their

---

3 On the company level that is. On the industry level it is well-established that social screens tend to favour growth-oriented and volatile stocks such as technology stocks (Bauer et al., 2005; Cerin & Dobers, 2001; Geurard, 1997; Kurz & DiBartolomeo, 1996; Vennix, Riegel, & Dunn, 2004). Statman (2000) found that the stocks on the US Domini Social 400 Index were slightly riskier than those on the S&P 500, and Vermier, van der Velde, & Corten (2005) found that the total risk of six European sustainability indices was higher than the risk of their conventional counterparts.
contribution to natural hazards that are not included in the rating instruments. This social trap character (cf. Platt, 1973) of company behaviour, where each company acts for its own advantage with a delayed negative outcome for the group or society at large, is not explicitly recognised in the ratings.

Assessment of Environmental and Social Performance

Financial Risks

Financial Outcomes

Environmental and Social Risks

Environmental and Social Outcomes

Figure 1. A conceptual model of corporate social performance analysis.

A final issue on validity is whether acting on the advice of SRI analysts enhances investors’ financial outcome or not. This is a matter that researchers have tried to elucidate (for a review, see Sparkes, 2002). Are investments in SRI funds more profitable than investments in non-SRI funds? Many times, researchers base their analyses on an already established categorization of funds into SRI and non-SRI. However, should the initial classification be invalid (Hoggett & Nahan, 2002), the endeavour is pointless. A better approach would be to establish independent measures of sustainability, based on diverse sets of environmental and social indicators, and relate these to financial outcomes.

Investment decisions/treatment

Diagnoses and advice is fed into the decision making of institutional investors. As indicated above, such advice can be presented in composite or multidimensional form. The form, in turn, has implications for negative versus positive screening.

To assess CSP is a multidimensional task (Carroll, 1979; Griffin & Mahon, 1997). Yet, investors sometimes adhere to a single criterion and apply a deontological rule for their decision making. They do not evaluate the financial, social or environmental consequences of their decisions. Rather, they decide in advance not to invest in any company that, for example, has business interests in nuclear weapons systems. Typically, negative screening of this type is applied with regard to conventional moral issues such as weapons, tobacco, pornography, and alcohol (Sparkes & Cowton, 2004). The outcome of this procedure is heavily dependent on which cut-off criteria are specified for particular company activities (Entine, 2003). Lenient criteria could result in an almost unaffected universe of companies to invest in, while
strict criteria may be seen as incurring financial risks which are too high, since they limit the diversification opportunities.

Negative screening sometimes serves as a first stage in the investment process. Companies that are not excluded are evaluated in a second stage by means of positive screening or a “best-in-class” method, where only those companies with the top CSP scores within each industry sector are available for inclusion in the investment portfolio. This leaves the institution with the decision to choose among these companies. At least two kinds of trade-offs emerge. When both social and environmental dimensions are relevant for the investment decision (and when ratings are presented multidimensionally), a balance must be struck between the two. In addition, the importance of sustainability aspects relative to the importance of more traditional financial aspects must be settled.

In case the analyst is responsible for striking a balance, a clear goal function and the weight attached to each dimension should preferably be communicated to the client. Assuming that the goal function targets financial risks, the importance of each dimension for predicting company or industry risk ought to be established: “For industry X, the social dimension is 1.7 times more important than the environmental dimension, and this differentiated weighting is supported by these data.” Admittedly, this is a recommendation that applies in an ideal world. Still, the message is that trade-offs should be made explicit and documented. Over time, this will facilitate consistency control. Alternatively, the assessment of the various criteria should be kept separately within each dimension. It is then up to the institutional investors to apply their own value function and make their trade-offs accordingly. The same advice regarding consistency control is appropriate.

Hence, when negative screening is applied, some companies are excluded. Which companies, and how many, is determined by the specific criteria and their cut-offs. When positive screening is used, the decision process is more complicated in that trade-offs have to be made.

Conclusion and recommendations

While the huge methodological problems associated with measuring CSP must be recognised, it is nonetheless for several reasons important that ratings of firms’ social and environmental performance are available within the investor community. Firstly, when extra-financial information about companies is translated into figures, it is more likely to be taken into account by institutional investors than if presented in qualitative form, as commonly is the case in company reports and other official documentation (Hassel, 2006; Guyatte, 2006). Secondly, ratings serve as a basis for sustainability indices, such as the FTSE4Good, Domini Social 400, and Dow Jones Sustainability indices. The mere existence of such indices is a potential driver of sustainability issues among companies aspiring for inclusion (Oulton, 2006). Thirdly, ratings facilitates for institutional investors to employ best-in-class or index-based SRI portfolio selection strategies (Sethi, 2005).

The usefulness of rating schemes for society, environment, and investors is however dependent on whether these schemes are based on the right criteria. While often being explicit about selecting criteria likely to affect financial risks and opportunities, SRI rating agencies generally refer to one or another UN convention or international agreement which they claim provide the foundations from which their criteria are developed. However, such conventions and agreements are often painted with a broad brush, and there is a significant lack of transparency as to how they are translated and operationalised by the analysts. Even though
SRI rating agencies generally present partial descriptions of their methodology, the exact operations between input (UN conventions etc.) and output (criteria/ratings) is still much of a black box.

Although broadly convergent, the criteria used in order to assess corporations’ social and environmental performance do, as we have shown, differ between rating agencies. There are several possible explanations to this. Criteria may be chosen or rejected on pragmatic grounds; for example, due to the ease with which information can be obtained. Having different criteria than other rating agencies may furthermore function as a branding exercise; a way to distinguish oneself from others, thus developing a competitive edge in the SRI analyst market. Both of these explanations imply that criteria are not selected purely out of scientific considerations. Still, disagreements could also be “real,” indicating that views diverge between SRI analysts as to what the relevant criteria for measuring CSP are. If so, this would be hardly surprising, as sustainable investment is still a fairly young area of research. Disagreement hence perhaps indicates that, at this point in time, knowledge about what criteria are most appropriate for assessing corporate social and environmental conduct is not well-advanced.

To the extent that rating agencies do rely on science when establishing CSP criteria, they may largely depend on having scientific alerts brought to their attention via media and NGOs. This is unlikely to be a proactive enough approach for institutional investors who want to make use of their ratings. While media hype can be beneficial if heightened interest in a problem area results in raised standards and a sharpening of the relevant CSP criteria on a permanent rather than a temporary basis, it is also conceivable that standards will be reverted once the collective memory of the investor community and the media fades. For example, while accounting transparency issues has gained greater weight in CPS ratings in the wake of the Enron and WorldCom scandals, next question-of-the-day threatens to overshadow this issue in the future. In order to reduce this risk, increased collaboration and cross-fertilization between analysts working within the SRI industry and scientists/academics would be beneficial. The scientific research community may not yet be at a point at which a complete list of CSP criteria could be fully justified (Griffin, 2000; Rowley & Berman, 2000), but greater transparency in the operations of SRI rating agencies (Cerin & Dobers, 2001), coupled with enhanced collaboration between analysts and scientists, would facilitate progress in the field. A list of CSP criteria produced through an interplay between natural scientists and SRI analysts appears more promising when it comes to predicting and evading environmental risks and their financial consequences than a list where criteria are – as suggested currently often is the case (Entine, 2003; Sethi, 2005) – to some extent selected arbitrarily or out of convenience.

Is it then possible to compose an accurate aggregate measure of CSP? Some authors suggest that CSP research should take other directions. Rowley and Berman (2000) argue that even if a linear relationship between CSP and CFP does exist, it is unlikely that these variables will be similarly related under all conditions. Instead of scanning data for statistically significant results linking CSP and CFP, researchers need to apply a contingency approach, examining the conditions under which a relationship (positive or negative) should be expected. Rowley and Berman warrant less focus on the CSP-CFP question and more focus on understanding phenomena (e.g., why stakeholders take action) that are part of the mechanisms creating relationships between types of social behaviour and other factors, including CFP. In a similar vein, Griffin (2000) suggests future CSP research to emphasize societal expectations, and the iterative process of firms’ responses to societal expectations.
Research questions should involve (a) defining relevant societal expectations, (b) devising mechanisms or processes to meet societal expectations, and (c) how societal expectations change over time. Hoffman and Bazerman (2006) point out that the present debate between sceptics and proponents of SRI has hardened into an ideological conflict between intractable positions, where one side claims that sustainability comes with a cost (a win-lose situation) and the other claims that sustainability is good for business (a win-win situation). Both positions are easily contradicted by the opposing side by means of singular examples of firms’ financial performance that seem to prove the contrary. The key to resolving this circular and contentious debate is, Hoffman and Bazerman (2006) argue, the recognition that corporate measures taken to promote sustainable development is sometimes profit-compatible and sometimes not. Moving beyond the simple question “does it pay to be sustainable,” this thinking leads to the question “how and when does it pay to be sustainable,” for specific companies and in specific circumstances (see also Howard-Grenville & Hoffman, 2003).

Possibly, a positive CSP-CFP link needs not to be established in order for institutional investors to take social and environmental issues seriously. UK evidence shows that some 42 percent of individuals who invest in SRI retail funds are willing to sacrifice some financial performance, and believe that they do so (Lewis & Mackenzie, 2000). Indeed, people are (to an extent) willing to pay taxes in order to benefit from a healthy society and environment. Furthermore, since pension funds have a distal time horizon for their investments, factors with no foreseeable tangible financial effects ought to be taken into consideration if likely to affect the society and environment in which the beneficiaries will be living their old age. Against this common-sense argument it is often objected that if institutional investors such as pension funds apply SRI selection criteria, they may fail their fiduciary responsibility not to compromise financial returns. Still, as long as SRI alternatives can demonstrate comparable financial performance to their conventional counterparts, this should not pose a legal obstacle for institutions to invest sustainably (Sethi, 2005). If a threshold proportion of institutional investors start taking social and environmental aspects into account – even though the link to financial performance has not been firmly established – SRI may gain momentum and others follow suit. A hurdle yet to overcome before SRI is likely to get into mainstream institutional investment processes proper is however that even sympathetic institutions tend to evaluate their fund managers’ performance on a quarterly, or half-yearly, basis in relation to an index benchmark, thereby discouraging pension funds from investing with an equally distal horizon as that of their beneficiaries (Guyatte, 2006; Mackenzie, 2006; Sethi, 2005). While this issue remains to be resolved, much is still to be done when it comes to improving the CSP measures currently in use within the investor community. With these developments well on their way, the case for institutional investors to integrate social and environmental considerations in their core investment analysis processes would be drastically enhanced.

References


APPENDIX

<table>
<thead>
<tr>
<th>Analyst agencies</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
</table>

BACKGROUND (CONVENTIONS REFERRED TO ETC.)

- UN conventions (human, child, ..)  X  X  X  X  X
- UN Global Compact  X  X
- ILO  X  X  X  X
- OECD  X  X  X
- Amnesty
- Rio declaration
- Weapons convention  X
- Environmental conventions  X
- Indigenous people  X
- Anti-personal landmines  X

INTERNAL ENVIRONMENTAL POLICY CRITERIA

- Formalised policy  X  X  X  X  X  X  X
- Management engagement  X  X  X  X  X
- Environmental management system  X  X  X  X  X
- Public reporting  X  X  X  X  X
- Dialogue  X  X  X  X
- Staff education  X

ENVIRONMENTAL PERFORMANCE CRITERIA

- Material (weight; % recycled)  X  X  X  X  X
- Energy  X  X  X  X  X
- Water  X  X  X
- Bio-diversity  X  X  X
- Emissions  X  X  X  X  X  X
- Waste  X  X  X  X  X
- Products  X  X  X  X
- Services  X  X
- Transport  X
- Demand suppliers
- Compliance with regulations  X  X
- Hazardous substances  X
## INTERNAL SOCIAL POLICY CRITERIA

<table>
<thead>
<tr>
<th>Strategy and Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formalised policy</td>
</tr>
<tr>
<td>Board structure and practice (independence, disclosure directors’ pay, ..)</td>
</tr>
<tr>
<td>Senior CSR officer</td>
</tr>
<tr>
<td>Equal opportunities</td>
</tr>
<tr>
<td>Trade unions treatment</td>
</tr>
<tr>
<td>Supply chain labour standards</td>
</tr>
<tr>
<td>Dialogue</td>
</tr>
<tr>
<td>Ethical code support (bribary, corruption,..)</td>
</tr>
<tr>
<td>Women on the board</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Human Capital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment stability</td>
</tr>
<tr>
<td>Staff education</td>
</tr>
<tr>
<td>Employment terms</td>
</tr>
<tr>
<td>Health and Safety policy/program</td>
</tr>
<tr>
<td>Compensation plans/payments</td>
</tr>
<tr>
<td>Employee satisfaction measure</td>
</tr>
</tbody>
</table>

## SOCIAL PERFORMANCE CRITERIA

| Donations/Charity                         | X | X | X |     |
| Community support programs                | X | X | X |     |
| Stakeholder engagement activities         |     | X | X |     |
| Product impact assessment                 |     | X | X |     |
| Developing country support program        |     | X |     |
| Awards received from external parties     |     | X |     |
| Fines and penalties assessed against the organization |     | X |     |
| Customer satisfaction measure             |     | X |     |
| Supplier                                 |     | X |     |
DATA COLLECTION AND ANALYSIS

<table>
<thead>
<tr>
<th>Data Collection</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questionnaire</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interviews company</td>
<td>X</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interviews others</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Official documents</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Others (NGOs, media)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data reporting</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Unidimensional</td>
<td>X</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multidimensional</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Differential weighting</td>
<td>X</td>
<td>X</td>
<td>n/a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal weighting</td>
<td>n/a</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sector/Industry sensitivity</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Country-specific</td>
<td>n/a</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Application</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative screening</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive screening</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Engagement</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sustainability indices</td>
<td>X</td>
<td>n/a</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>