

## **An Analysis of the Methodologies adopted by CSR Rating Agencies**

### **Abstract**

**Purpose:** This paper analyses the similarities and differences in the methodologies adopted by CSR (corporate social responsibility) rating agencies.

**Design/methodology:** We gather secondary and primary evidence of practices from selected agencies on the methodologies and criteria they rely upon to assess a firm's CSR performance.

**Findings:** We find evidence of similarities in the methodologies adopted by the CSR rating agencies (e.g. the use of environment, social and governance themes, exclusion criteria, adoption of positive criteria, client/'customised' input, quantification), but also several elements of differences emerge, namely in terms of the thresholds for exclusion, transparent vs. confidential approach, industry-specific ratings, and weights for each dimension. Drawing from Sandberg et al.'s (2009) conceptualisations, we tentatively argue that this mixed picture may reflect competing organisational pressures to adopt a differentiation approach at the *strategic* and *practical* levels whilst recognising, and incorporating, the 'globalising' tendencies of the CSR business at the *terminological* levels.

**Implications:** Although our data is based on a relatively small number of agencies, our findings and analysis convey some implications for users of CSR ratings and policy-makers; particularly in light of the recent Paris 2016 Agreement on Climate Change and the increased emphasis on the monitoring of social, environmental and governance performance.

**Originality:** We contribute to the literature by highlighting how key intermediate rating organisations operationalise notions of CSR.

**Keywords:** Corporate social responsibility (CSR); CSR rating agencies; Socially Responsible Investment (SRI)

**Classification:** Research paper

## **1. Introduction**

There has been increasing attention paid by market and societal actors to the social performance of corporations and the gradual embedding of ethical investment in mainstream finance. As a result, there has been a noticeable reliance on CSR (corporate social responsibility) Rating Agencies as intermediate institutions which claim to provide a structured and reliable assessment of an enterprise's engagement with the multiple, and often ambiguous, facets of corporate social responsibility (Chelli & Gendron, 2013; Archel et al., 2011; Scalet & Kelly, 2010; Bessire & Onnee, 2010).

CSR rating agencies aspire to bring a sense of order to the field (Scalet & Kelly, 2010) by collating insights from different information sources (including from the reporting entity itself) and presenting generally quantifiable indicators of corporate social performance (Chelli & Gendron, 2013; Malsch, 2013). In turn, these indicators inform decision makers on an investment's social credentials (e.g. ethical investment funds, pension funds) and the risks attached to an enterprise's ethical profile (Sandberg et al., 2009; Scalet & Kelly, 2010; Bessire & Onnee, 2010). The users of ratings agencies also include non-financial stakeholders such as government (e.g. in procurement; refer to D'Hollander and Marx, 2014), non-governmental organisations (NGOs) and other user or pressure groups. However, financial stakeholders and investment decisions tend to be the main motivations for the multiplication and increased visibility of agencies worldwide (Rhodes & Soobaroyen, 2010; Hedesstrom et al., 2011; Sustainability, 2012; Malsch, 2013; Scott et al., 2014; Dilla et al., 2016). At the same time, there remains a 'black box' in terms of understanding how ratings are devised by agencies and subsequently relied upon by investors (Dilla et al., 2016).

Similar to the case of other accreditation institutions (e.g. credit rating agencies; quality accreditation firms) and in tandem with accounting firms and professional bodies (Malsch, 2013), CSR rating agencies deploy indicators and other ranking criteria which enables them to provide their services to market players and to represent themselves as credible assessors of corporate social performance. This credibility is also visibly reflected in an increased reliance on the data provided by CSR rating agencies (principally categorised in terms of environmental-, social- and governance-related items) by many academic studies. This is motivated by a search for quantitative proxies of corporate social performance (Igalens et al., 2008; Bessire & Onnee, 2010; Semenova & Hassel, 2014; Scott et al., 2014), particularly those deemed to be outside the direct control of the reporting enterprise (Parguel et al., 2011). Bessire & Onnee (2010) argue that many agencies adopt a utilitarian ideology privileging scientific methods, quantitative measures (metrics), financial-led clients and transparency of the methodologies. However, anecdotal evidence suggests that many agencies are not very transparent (SustainAbility, 2012) and noted differences have emerged between the ratings themselves (Woods & Jones, 1995; Entine, 2003; Koehn & Ueng, 2005; Chatterji & Levine, 2008; Sandberg et al., 2009; Hedesstrom et al., 2011). So far, there have been limited assessments of their methodologies (Hedesstrom et al., 2011; Escrig-Olmedo et al., 2010).

Informed by the above, this study examines the ratings methodologies of a group of CSR rating agencies, focusing on the following key research question: *How do these agencies operationalise CSR in their evaluation criteria and methodologies? What are the key similarities and differences?* To answer these research questions, we first carried out a secondary data analysis of the rating methodologies as disclosed on the agencies' websites, including publicly available documents relating to their services and activities. Secondly, we

contacted a group of agencies to examine their criteria, CSR dimensions and the extent to which industry differences are accounted for.

Briefly, the findings, albeit from a relatively small number of organisations, reveal that the ratings methodologies adopted by the responding agencies display some similarities in relation to the CSR themes being assessed, the exclusion criteria, the adoption of positive criteria and the prominence of client-led criteria. At the same time, areas of differences are observed in terms of the threshold for exclusion, the use of industry-specific criteria and weights to assess social performance.

Our paper contributes to the literature as follows. Firstly, our findings extend the prior insights from Hedesstrom et al. (2011) and to a lesser extent, Escrig-Olmedo et al. (2010), in that the rating methodologies reflect a rather 'adaptive' attitude to the measurement of environmental, social and governance performance, in the sense that the methodologies change from one CSR rating agency to another. Exclusion criteria are in practice not particularly restrictive and what could be considered as a 'material' involvement in controversial businesses appears to vary significantly. Secondly, Igalens et al. (2008) and Sandberg et al. (2009) highlighted the potentially national or cultural character ingrained in some ratings (e.g. is gambling an acceptable business in given regions/countries?). Whilst acknowledging that we can only draw our conclusions from a limited set of countries and these will need further validation, we note that such national or cultural aspects did not emerge from our data. Thirdly, Chelli & Gendron (2013) investigated 37 rating agencies worldwide but were more concerned with the *commonality* of their public discourses and the quantification agenda underlying the methodologies adopted by ratings agencies. However, when these agencies were directly prompted for information, we instead noted an apparent

combination of *commonality* and *difference* of the methodologies and criteria, principally arising due to market and commercial rationales. In our overall discussion of the findings, we draw upon Sandberg et al. (2009)'s conceptualisations to provide a tentative analysis of the findings. Admittedly, an assessment of a firm's corporate social performance on the basis of a given rating which then differs from the one devised by another rating agency may be acceptable for some informed users (e.g. analysts; investors), but not necessarily for others (e.g. social stakeholders). Nonetheless, our study has implications for all users of CSR ratings in that it highlights where areas of differences between agency ratings are more likely to emerge and thereby enable users to critically evaluate the reliability and relevance of these criteria.

The remainder of the paper is structured as follows. The next section briefly reviews the extant literature and is then followed by an explanation of the theoretical perspective. Subsequently, the data and methods are outlined and the findings and analysis are presented. Finally, the implications and contributions are discussed in the concluding section of the paper.

## **2. Literature Review**

Several studies have used or exploited ratings, scores and rankings produced by so-called sustainable indices, CSR rating agencies or even CSR information providers but few of them have questioned the validity and reliability of these measurements. A notable exception is the work published in a recent special issue on the reliability of sustainability indicators (refer to Scott et al., 2014). For instance, Allouche & Laroche (2006), Griffin & Mahon (1997) and others (refer to Margolis & Walsh, 2003; Orlitzky et al., 2003 for a more extensive list of such studies) relied on CSR ratings to examine the relationship between the corporate

social performance (CSP) and the financial performance of the firm. A number of studies also attempted to propose different approaches for measuring CSR (Dillenburg et al., 2003; Gauthier, 2005; Knoepfel, 2001; Knox & Maklan, 2004; Marquez & Fombrun, 2005; Mattingly & Berman, 2006; Miller et al., 2007). It should be noted that the literature does rely on a plethora of terminologies in relation CSR ratings, such as Socially Responsible Investment (SRI) metrics, Sustainability Ratings, Economic, Social and Governance (ESG) measures. In this paper, we have selected the term ‘CSR ratings’ and consider it to be largely interchangeable to the other terms for the benefit of the discussion, although we do acknowledge that there are differences/overlaps between these terms.

Some authors critically assessed the quality of ratings of the most used or known agencies or indices. Igalens & Gond (2005) have conducted an analysis of the data provided in 2000 by ARESE, the leading French rating agency at the time, while Sharfman (1996), Mattingly & Berman (2006) and Chatterji et al. (2009) have assessed in different ways the most widely used ratings, namely those by Kinder, Lydenberg, Domini Research and Analytics (KLD). In spite of their increasing popularity, CSR ratings have been rarely evaluated and have been criticized for their own lack of transparency (Chatterji et al, 2008). Hedesström et al. (2011) consider that CSR ratings agencies should communicate to investors and companies the theoretical and empirical basis upon which environmental evaluation criteria are selected, and address the criticisms relating to the validity of CSR ratings. The issue of insufficient empirical and theoretical justification has also been previously discussed. Entine (2003), for example, argues that raters assign high marks to firms that are later more likely to be embroiled in scandals. Hawken (2004, p. 16) criticised socially responsible investing (SRI), noting that “*the screening methodologies and exceptions employed by most SRI mutual funds allow practically any publicly-held corporation to be considered as an SRI portfolio*

*company.*” Indeed, there does not seem to be any set uniform standards which all CSR rating agencies could comply with. At the same time, one may question whether it is desirable to have uniformity in the assessment of ‘what is (or ought to be)’ corporate social performance (Sandberg et al., 2009). Not only do CSR rating agencies appear to differ about how to weigh a given criterion (such as the treatment of employees versus the respect of the environment), but they also differ about what constitutes relevant data for assessing the criteria. Thus, it would be impossible to compare results across agencies (Scalet and Kelly, 2009), leading to some academic work examining the fit between competing ratings (e.g. Sharfman, 1996; Chatterji et al., 2008; Semenova & Hassel, 2014). This also raises the question of consistency because when respective rating schemes of the different analyst organizations vary in terms of evaluation criteria, this could result in the same company receiving favourable ratings from some analyst organizations and comparatively poorer ratings from others (Hedesström et al., 2011). Dilla et al. (2016) equally mentioned this issue when they examined how nonprofessional investors use their SRI screens and found that nonprofessional SRI investors relied on filtered environmental performance information from rating agencies to make comparisons across individual companies. While admitting that the filtered information from rating agencies may make it easier for investors to compare the social performance of individual companies, the authors raise the issue of a lack of consistency due to the presence of substantial differences in the type of environmental performance measures reported by those rating agencies.

This lack of consistency and transparency may be troubling to many, particularly for those who expect that CSR ratings can be objectified and made functional, thereby conveying a generally accepted reality of ‘corporate social performance’; akin to a bottom-line accounting metric. However, corporations are given mixed messages about what is being measured

(Dillenburg et al., 2003), with the question of ratings convergence having different implications for institutional investors, the evaluated firms, civil society and for academic research in the field (Hedesström et al., 2011). There has been very little empirical substantiation in the academic literature to underpin a debate about the need for ratings to be converging (Hedesström et al., 2011). For instance, the most examined CSR ratings were those by KLD. Sharfman (1996) studied the validity of the KLD data by comparing it to other more accepted measures of CSP. He found that KLD ratings correlated sufficiently with other measures of CSP thus indicating a sufficient level of reliability. Mattingly and Berman (2006) also assessed KLD ratings and found that positive and negative social actions are both empirically and conceptually distinct constructs and should not be combined in future research. Chatterji et al. (2009) find the 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. More recently, Semenova & Hassel (2014) examined the correlation between the different environmental scores assigned to companies by several CSR rating agencies and found some evidence of consistency between the different ratings, albeit this is limited to 3 agencies (KLD, ASSET4 and GES) and to US companies only.

Using data from several CSR ratings agencies such as KLD, Calvert, FTSE4Good, DJSI, and Innovest, Chatterji & Levine (2008) found that the major ratings did have a fairly low correlation with each other, supporting a market-driven differentiation strategy. Chatterji & Levine (2006) found similar results when they examined three major indices of the United-States (KLD’s Domini 400, Dow Jones Sustainability Indexes (DJSI), and the FTSE4Good), namely that the measurement systems, the resulting selections and rankings of companies



differed considerably. They reported several critical issues about the weighing systems, data collection, transparency and level of performance.

In terms of the weighing attached to various components of non-financial performance, Chatterji & Levine (2006) noted that the above-mentioned firms have quite different methodologies, which results in an unreliable comparison of social performance between companies. This lack of uniformity undercuts the ability of stakeholders to discriminate among more or less reliable ratings and allows corporations to pick and choose among rating outcomes so as to avoid pressure to address poor social performance. Finally, Chatterji & Levine (2006) pointed out a lack of transparency noting that there might be good reasons to have different weighing systems to use as benchmarks against a user's own portfolio, but the implicit motives which drive these differences in measurement are not elucidated by ratings firms. Indeed, a large amount of information remains impossible to obtain. This lack of transparency makes it difficult to understand what CSR ratings agencies are measuring and whether one can make reasonable comparisons between them (Chatterji & Levine, 2006). In a recent study Escrig-Olmedo et al. (2010) again highlighted the problem of consistency and found that there is no standard methodology for the evaluation of corporate social performance. The lack of consistency between different ratings creates an issue of comparability and the problem seems to be compounded when the positioning of a company and its CSR status are expressed symbolically, or when using a numerical scale (Escrig-Olmedo et al., 2010).

Despite these differences, most agencies seek to rate companies through a combination of positive and negative criteria; and these are based on an adapted Stakeholder Model (Fassin, 2009; Podnar & Jancic, 2006) and a series of global standards. Escrig-Olmedo et al. (2010)

also noted that companies are faced with a lack of information that makes it difficult for them to discover which actions would enable them to perform better (and be included in a CSR index) and investors equally face difficulties selecting sustainable companies for investment. More recently, Hedesström et al., (2011, p. 130) analysed the environmental evaluation criteria and ratings of seven European and North American agencies and found that ratings differed, depending on the rating scheme used. For example, comparisons of the environmental ratings for the automobile sector show that there is a fairly broad consensus across raters as to which car manufacturers have the worst environmental performance. However, there is considerable disagreement as to which ones are the best performers.

In conclusion to this section, there has been much concern expressed on the screening of investments deemed to be socially responsible or sustainable and although an assessment of the methodologies and criteria used by the CSR ratings have been singled out in a few studies, there remains a dearth of research on the rating agencies themselves.

### **3. Data and Methods**

In line with Chelli & Gendron (2013) and Hedesström et al. (2011), we relied on the guide published by the French Observatory of CSR (ORSE, 2012<sup>1</sup>) identify the current market players in CSR ratings. We accessed the latest version of this guide and combined it with other different sources to identify a list of 28 agencies (See Appendix 1). However, unlike Escrig-Olmedo et al. (2010) who also studied CSR information providers supplying sustainability indices, we only selected those agencies that produce and sell CSR ratings to companies and investors. A large majority of these CSR rating agencies was created before 2004 (21 agencies), while only 7 agencies were founded between 2007 and 2011. Most of them are European (20 agencies) while 5 agencies are from Asia, 2 from Australia and one

from Central America (Mexico). When analyzing the websites of these agencies, we noticed that six of them (CAER, IMUG, Greeneye, Ecovalores, KO-CSR and ECODES) are partners of the British agency 'EIRIS' and hence part of its network i.e. they apply the same methodology of rating as EIRIS<sup>ii</sup>. In addition, 'Ethibel', the Belgian agency was absorbed by its French counterpart 'Vigeo' in 2005. Thus, if we consider not the number of agencies but their approaches to assessment and evaluation that are supposed to be different, we concluded that we have 21 potentially different methodologies of ratings.

From an initial assessment of the public documents available on the websites of each agency, we noted a lack of information about their methodologies and even less so when it came to their evaluation criteria. This was raised by some authors like Escrig-Olmedo et al. (2010, p.448) who consider this lack of information as a result in itself. The only exception was *SAM Research*, a Swiss rating agency which publishes a guide on its website detailing its approach and methodology, which provided concrete examples of its evaluation criteria and the weights it assigns to them on the basis of the industry being evaluated. We developed a set of thematic questions (see Appendix 2) based on what agencies concretely do and we were particularly drawn by the presentation and classification made by SAM Research. These questions were also adapted from those used in prior work (Escrig-Olmedo et al., 2010; Hedesstrom et al., 2011; Sustainability, 2012). The first set of questions examines the exclusion criteria used by the agency. Considering that some approaches/methods of rating exclude certain industries for religious, moral or political reasons or for their involvement in activities that threaten the environment, the human health or the ecosystem, we asked the agencies if they used exclusionary criteria of any nature in their assessment process. Agencies also had to mention at which threshold of involvement they would consider the industry (or firm) to be excluded.

The second part of the questions sought to identify the CSR themes used by the agency in its assessment process. According to the most common definitions of CSR<sup>iii</sup>, a firm's responsibility is threefold: economic, social and environmental. So we asked the agencies about the CSR themes they consider in their evaluation process, in order to know if they use all three dimensions or only some of them. We also asked the agencies to supply the weight they assign to each dimension in their evaluation, regardless of the industry examined. Again, we aimed to assess if the evaluation system gives some preference to a specific dimension rather than another. For example, one CSR rating agency could focus on the environmental dimension either by measuring this theme only or by giving it a higher weighing compared to the other dimensions.

The third and final part looked at the selection criteria. In order to evaluate all the CSR dimensions/themes, rating agencies define, examine and check numerous criteria for each dimension. These criteria can be common to all the companies regardless of the industry to which they belong (general criteria) or can change according to the industry to which the assessed company belongs (industry-specific criteria)<sup>iv</sup>. We were not interested in the criteria themselves since they can vary from one agency to another depending on the qualification and the interpretation the agency makes of the dimension it evaluates and the priorities it identifies when assessing it. Instead, we preferred to focus on the nature of these criteria rather than their number, taking into account the large number of criteria that can exist for each dimension and the aversion of the agencies to display them. Thus, we asked our agencies if they use a general criterion when assessing companies or did they adapt their criteria according to the industry to which the assessed company belongs. A combination of both types of criteria was also possible. CSR rating agencies can employ general criteria and at the same time use some industry-specific criteria. In every case, it would be interesting to know

the weight of each type of criteria in the evaluation process. To contextualize the answers, we asked each agency to inform us of the typical weight of the general criterion it uses in the evaluation of each dimension. Similar to Hedesström et al. (2011) who focused on two industries (automobile and paper/forestry) to study environmental performance ratings, we focused on three industries (Banks, Electricity and Pharmaceutical companies) and asked for the weight of the industry-specific criteria. We have selected them as illustrations for our study but admittedly, a selection of different industries may lead to more different or similar assessments.

We made contact with the 28 CSR rating agencies by email to request for their participation. After further direct contacts, 10 agencies agreed to respond to our request for information underlying the prior difficulties encountered by researchers in the field (Escrig-Olmedo et al., 2010; Hedesstrom et al., 2011). Whilst our intention is not to claim a statistical generalization, it is also important to note that we made contact with the relevant specialists or relatively senior individuals in the firm (e.g. Research Director, ESG-Analyst, Head of sustainability services, Head of the agency (CEO, Director), Communications coordinator, and Head of market relations) which can ensure some degree of relevance in the responses. We conducted an average of 30 minutes to one-hour phone or Skype interviews with each of our respondents and where required, asked for clarifications by email. For the remainder of this paper and due to the confidentiality arrangements, we have coded each organization from RA1 to RA10.

#### **4. Findings**

##### *The exclusionary approach*

Table 1 below provides a summary of the exclusionary criteria as revealed by the responding agencies since a tabular format can provide a snapshot of the key insights from our

informants. Our enquiries revealed that 5 out of 10 rating agencies adopted an exclusion criterion to avoid firms that are connected to so-called controversial activities, whereas the other agencies do not have any explicit criteria. This does not however mean the agencies systematically apply the criteria on companies they are required to rate. Some informants indeed then explained that they examine the involvement of the companies they evaluate in controversial sectors and let their clients (investors) decide of the exclusion (or not) of the firms incriminated. This means the client/investor sets him/her-self the exclusion criteria the rating agency should apply to create the portfolio either by choosing from a list of exclusion criteria or by asking for specific ones. As stated by one manager:

*“We apply a customized approach which is supposed to suit the needs of investors and reflect their specific perspective on sustainability, hence no general answer is possible in this case as it varies from client to client” (RA3)*

This combined approach appears to be gaining ground, and thus limiting the exclusion criteria approach to the gathering of information about the involvement of the audited company in controversial issues; the final decision of excluding or including the company in the portfolio thereby resting with the client (investor). For example, the RA4 informant pointed out that in addition to the exclusions they operate on a standard basis, they also tracked data on animal testing and stem cell research:

*“We exclude Alcohol, Contraceptives, Tobacco, Gambling, GMO Food Production. Military/Army, Nuclear Power, Pornography are excluded from [our] indices on a standard basis, in addition to this we do track also Animal Testing, Stem Cells. The analysis takes into consideration also the nature of the involvement of a company (e.g. Military/Army includes: Cluster munitions and antipersonnel landmines; Nuclear, biological and chemical weapons; producers of conventional armaments etc.)” (RA4)*

## INSERT TABLE 1 ABOUT HERE

The exclusion criteria used by our agencies are in most cases billed as ‘ethical’ or ‘religious’ in nature (for 4 agencies out of 5). This could be explained by the religious origins of SRI and the tradition of avoiding “sinful products”, which can be associated to the development of the movement in the UK and Europe (Sandberg et al., 2009), as highlighted by the following website disclosure by a European CSR rating agency:

*“..tobacco, alcohol, gambling, GMOs, weapons, nuclear and adult entertainment. Controversial practices screening includes human rights abuse, corruption and damage to the environment among others”. (RAI0)*

Activities relating to the damage of the environment were also employed by two agencies (out of 5) as an exclusion criterion. As for the excluded industries, the vast majority of our agencies purported to avoid firms involved in weapon or arms manufacturing. Nuclear power and tobacco were also cited by 6 of those rating agencies but only 4 excluded the alcohol industry. To a lesser extent, other industries such as gambling, genetically modified food production (GMO), contraceptives, financial services, hotel industry, were also avoided. These examples again underline the focus made on ethical/religious and environmental issues. But the exclusion may also be normative and not only industry-related since some agencies stated that they would avoid firms that would rely on child labor for example, thereby being more in line with an acknowledgement of human rights and the unfair treatment of minorities or marginalized communities (Sandberg et al., 2009). Compared to the findings by Escrig-Olmedo et al. (2010), our respondents seem to have defined similar items for exclusion. What is however new from our research is the threshold or degree of involvement that would be deemed ‘unacceptable’.

The definition of an exclusion criterion is indeed not sufficient since CSR rating agencies also establish the degree of involvement of the firm in the prohibited industries, products or activities. From the comments of the informants, the threshold of involvement in controversial issues often depended on clients and varied from 5% to 20% (of sales revenue) for some of the agencies. For some others however, this threshold could be anything from 5% to 100% of involvement while another argued for the exclusion of a firm at any level of involvement in so-called controversial activities. For example, RA4's (a European agency) threshold is up to 5% for the most of controversial industries like alcohol, tobacco or gambling but for some other categories like cluster Bombs, anti-personal landmines, nuclear, contraceptives, pornography or GMO, the company will be excluded, irrespective of its level of involvement. Once again, some rating agencies insisted on the fact that it is to the client/investor to decide and that their role only consists in providing the maximum information about the companies they were tasked to rate:

*“We just provide exclusion criteria and companies' qualification to clients. Also we evaluate relative sustainability performance of the companies in the industries with exclusion criteria. The clients decide whether they will exclude it or not” (RA8)*

#### *The CSR themes*

With regards to the CSR themes, all of them highlighted the use of the social dimension, even if some went further by referring to labour practices, human rights or human resource management. The environmental dimension was mentioned by 9 agencies while only one agency has cited the economic dimension of CSR as one of the themes it evaluates (RA7). This reflects the focus on non-economic performance and in doing so, CSR rating agencies maintain a position that is complementary to financial rating agencies, since it is expected that



investors would gather information from both agencies to create their ethical portfolio. Moreover, 7 agencies out of 10 stated that they evaluate the corporate governance (CG) of the firms. Thus, there was a relatively high level of commonality in terms of each theme.

Furthermore, as reported in Table 2, the weight for each CSR theme in the process of evaluation was not revealed by half of our informants, either arguing that such information is confidential or that it is customized and thus depends on the client. Whilst this could be interpreted as a lack of transparency, at the same time such a decision may be motivated by a strategic choice; because by doing so CSR rating agencies convey the message that they possess a specific know-how (i.e. a form of intellectual property) and/or an ability to adapt to customer/investor needs. As an example, one agency manager seems to rather boast of the breadth of coverage and complexity of the weights/models underlying the ratings while another informant put forwards the firm's technical expertise to justify this confidentiality:

*“We cover more than 150 industries sectors. Each of them have specific weighting configuration following sector analysis on materiality. Unfortunately, we cannot disclose this information” (RA1)*

*“The weightings are not fixed but adjusted to each sector, country and company. And even in the case of each company the weightings are transformed through a mathematical formula. Our system quantifies the value of a companies' management & business processes. We seek out the value drivers in each company” (RA6)*

**INSERT TABLE 2 ABOUT HERE**

For the other respondents who did reveal the weight of each CSR theme, environmental and social themes always predominated in terms of the weights assigned in the evaluation process but their weight varied according both to the rating agencies and to the industry they examined. For example, the environmental theme was weighed the highest in the electricity sector while the social theme was weighted highest in banking and pharmaceuticals. The third component - CG - was the less weighed one when mentioned, varying from 8% to 39% in all industries reviewed by the respondents. We however could not draw out a cogent explanation to the differences in weighting among rating agencies given the inconsistencies and somewhat contradictory assessments. For example, the environment was weighted from 36% to 40% by RA2, RA9 and RA10, in the pharmaceutical industry while RA7 weighted it only at 17% in the same industry. The same observation was noted for the social theme which was weighted from 21% to 62% in the banking industry, from 30% to 83% in the pharmaceutical industry and from 22% to 46% in the electricity sector (See Table 3). In each case, the difference between the maximum weight and the minimum weight might appear problematic but not so to the respondents. To justify such differences, one informant explained that his firm will:

*“apply a customized approach and it very much depends on the weightings, criteria and investment policy each single investor chooses to apply. Therefore, there are no general answers to the question you have asked but rather varying forms and contexts” (RA3)*

Hedesstrom et al.'s. (2011) study of environmental ratings in two industries (automobile and paper/forestry) did show that CSR rating agencies significantly disagreed on which firm was the best performer in each sector. Contrastingly, our findings actually reveal large differences in the weights attached to each dimension within the same sector and thereby explaining (arguably not to a full extent) the different outcomes by individual rating agencies.

*“No, these three dimensions are not weighed the same. Moreover, analysts have the possibility to change such weights according to companies’ specificities” (RA10)*

The issue seems more pronounced in the case of the social dimension and this potentially reflected very different conceptualizations of ‘corporate social contributions’ by the CSR rating agencies; an insight which has not been previously uncovered by Escrig-Olmedo et al. (2010) and Hedesstrom et al. (2011).

#### *The typology and weight of selection criteria*

All the informants acknowledged the use of industry-specific criteria to assess the social performance of the firms they evaluate. Yet most of them (8 out of 10 agencies) relied on a general criteria applied to all industries. Not all the agencies agreed to reveal the weights assigned to general or industry-specific criteria in the evaluation of each CSR dimension/theme; two agencies did not give the weight of their general or industry-specific criteria again arguing it is either confidential or customized. Also, two other agencies which only used industry-specific criteria did not provide their weight on the same grounds. The confidentiality of the information was justified by our respondents on the need to preserve their expertise and know-how. For example, RA8 declared: *“The weight of general criteria in each CSR theme is one of core knowhow of ESG research and confidential”*. Also RA1 informant said it has 21 CSR criteria for 150 industries but declined to provide information about them:

*“We have 21 CSR criteria which we apply to 150 industry sectors. Activation and weighting of criteria is performed based upon materiality analysis. We are not authorized to disclose activation and weighting of criteria for each of the industry sectors”*.

In several of the replies to our questions, the mentioning of a 'customized' weight appeared to be another way to avoid disclosing data. Whilst this may also mean that CSR rating agencies adapt the weight of their criteria to the needs of the client/investor, this in practical terms could not be done without the expertise of the agencies and their advice. So in both cases, rating agencies sought to put forward their know-how and yet at the same time, avoid an assessment as to whether the criteria and weights can be deemed appropriate. Some agencies do mention this aspect on their website, such as follows:

*“In order to be able to analyse comprehensively the diverse environmental and social challenges relating to the activities of companies, [RA2] has developed a pool of indicators. These currently number approximately 700. For each company, an average of 100 indicators is selected from this pool on an industry-specific basis so that a targeted evaluation of the problems specific to that company can be carried out.” (RA2)*

These points suggest there has been a growing trend to incorporate more ambiguity in the techniques of socially responsible investment (SRI) with a view to recognise the different values and priorities of companies and investors. Hence, it is likely that rating agencies are adopting what we refer to as an adaptive approach by responding to client demands and expectations. Hence, what appears to be a 'problematic' heterogeneity from an academic or outsider perspective of transparency and accountability may be less so in the eyes of investors and clients<sup>v</sup>. This heterogeneity can also be observed in the way agencies apply a weight to the general or industry-specific criteria. Indeed, this weight may be the same for all the dimensions/themes of CSR evaluated or vary from a dimension to another depending on rating agencies (See Table 3 below). For example, RA4, RA6, RA7 and RA10 claim they applied the same weight to their general criteria regardless of the theme evaluated (and of course, as a consequence, the same weight to the industry-specific criteria). Yet, this weight

differed subsequently between them as we record 100% for the first agency, 67% (2/3) for the second, 40-50% for the third and 70-90% for the last one. The second group of agencies applies different weights to its general criteria depending on the CSR theme being evaluated. For example, RA2's weight for general 'social dimension' criteria represents between 60% and 70% while it fluctuates from 40% to 50% for RA9. The general criteria for the environment are weighted 40-60% by RA2 and 30-50% by RA9 which otherwise uses only general criteria to evaluate the corporate governance dimensions.

**INSERT TABLE 3 ABOUT HERE**

The same observation could be made for the industry-specific criteria. If we admit that the weight varies depending on the industry evaluated, no clear explanation could be drawn to justify the differences in weight between the different agencies. For example, when the rated company belongs to the banking sector, the environmental performance is evaluated on the basis of 50% of industry-specific criteria by RA2 and RA9 but the same figure comes to 70% weight by RA7. In the case of the electricity sector, this weight is once again the same for RA2 and RA9 (i.e. 60%) while it rises to 83% for RA7. Another example is that of the social dimension which is evaluated on the basis of 30% and 35% of industry-specific criteria respectively by RA2 and RA7 in the same sector. RA9 uses however 70% of industry-specific criteria to evaluate this dimension in the same sector. Evidence of further differences between these agencies is observed when the rated company belongs to Pharmaceuticals with 40%, 56% and 70% of industry-specific criteria respectively for RA2, RA7 and RA9. A final example which shows once again the large diversity of methods among CSR rating agencies is that of RA4 which applies only industry-specific criteria for two of its CSR dimensions (products and production process) and only general criteria for the other CSR dimensions,

namely environmental strategy and policy, environmental management, employees and human capital, community relations, markets, corporate governance and shareholders.

## **5. Discussion and analysis of findings**

Overall, the findings from our sample of CSR rating agencies point to a picture of both similarities and differences in terms of ratings methodology, and are to some extent consistent with the findings of Escrig-Olmedo et al. (2010) and Hedesstrom et al. (2011). In an attempt to provide a tentative explanation of the findings and potential reasons for these similarities and differences, we dwell further by relying on the conceptual insights suggested by Sandberg et al. (2009).

Sandberg et al. (2009) emphasises the heterogeneity of the SRI movement and assert that there are four levels or dimensions where differences (and similarities) can exist, namely the *definitional*, *terminological*, *strategic* and *practical* levels. The *definitional* level relates to how SRI is defined within the field or by the CSR rating agencies (e.g. what is socially responsible investing? What is corporate social responsibility? What are environmental, social or governance aspects?). Sandberg et al. (2009) argues that there is overall broad agreement between the key actors (investors, analysts, rating agencies) on the definitional facets of CSR. However, diversity is said to exist on aspects relating to *terminological* (i.e. use of terms such as governance, social, environmental), *strategic* and *practical* levels. In relation to the strategic level, and drawing from Sandberg et al. (2009), it can be argued that CSR rating agencies may choose to privilege different approaches when advising clients on a company's level of social performance. For example, it can emphasise a negative (what to exclude) or positive (what to include) criteria about company's sustainable activities. Finally, the practical level relates to the actual implementation of the detailed criteria by different organisations and

where the potential for differences is at its highest; arguably a key level for assessing the more detailed aspects of the methodologies adopted by rating agencies.

From the responses we obtained on the type of criteria used by the ratings agencies, we note that many definitional, terminological, strategic and practical aspects contained in the rating methodologies are quite similar and are reflected by the use of common terminologies, namely the CSR dimensions (Environmental Social and Governance) and exclusionary themes, the adoption of positive criteria and client/'customised' input and perhaps unsurprisingly, a homogeneous reliance on seeking a quantification of these dimensions. We also note a picture of homogeneity which appears to go beyond the definitional and terminological levels as initially suggested by Sandberg et al. (2009), namely in relation to exclusionary themes, positive criteria and the prevalence of customer/client input (strategic level).

At the same time, there are significant variations at the strategic and practical levels namely thresholds for exclusion, the highlighting of key sub-dimensions, the extent of transparency or confidentiality, the industry-specific ratings, and the weights for each dimension. From a research perspective, the differences can still be characterised as being rather problematic for CSR researchers, stakeholders and other external users who may rely on ratings to provide a systematic, consistent and credible assessment of the social performance of a company. Notwithstanding, we partially concur with Sandberg et al.'s (2009) comments that the heterogeneity of SRI appears to persist at the strategic level (e.g. confidential vs. transparent approach; industry vs. general ratings) but more predominantly at the practical level (e.g. thresholds, weights).

In terms of the likely explanations for such differences, Sandberg et al. (2009) highlight three key reasons. First, differences may arise due to cultural factors relevant to particular regions or nations, and how concerns about CSR might have emerged; for example, in the UK compared to the US. In a nutshell, development in the former is typically associated to religious and moral imperatives whilst is concerned about diversity and human rights traditionally dominated in the latter. The second reason for differences in ratings methodologies can be related to the values, norms and ideology of the different actors involved (Sandberg et al., 2009) and rating agencies having to respond to varying expectations by clients, regulators, media and the general public amongst others. This can include changes within the ratings industry as a result of mergers, acquisitions or the involvement of larger international firms. Hence, this leads to the view that “...*understanding SRI heterogeneity from the perspective of different stakeholders suggests that differences in how SRI is understood or defined could depend on who you are and what objectives you have.*” (p. 525).

A third reason which may explain differences in ratings methodologies is the need for commercial players, such as CSR rating agencies, to achieve and maintain a market position, and this implies the need to engage in a differentiation or segmentation strategy (Chatterji & Levine, 2008). In a similar vein, O'Rourke's (2003) argues that CSR rating agencies have to develop a strong identity to attract investors. This can imply that a rating agency will adopt various tactics, whether of a linguistic, discursive and/or technical nature, in the design of its ratings methodologies to reinforce its market position. Also, Sandberg et al. (2009) highlight that there may be little incentive for entities involved in an SRI environment (such as CSR rating agencies) to engage in a homogenisation of ratings as long as market participants (i.e.



client-investors) do not demand for some form of common conceptualisation and metrics to operationalize CSR.

In consideration of the factors which could explain similarities and differences (as put forward by Sandberg et al., 2009), we find very little in the way of national culture or norms which might underpin the observed differences and similarities. Instead we contend that the differences in methodologies seem to be primarily associated to a market-led strategy of differentiation, which incentivises heterogeneity and distinctiveness at the strategic and practical levels; thereby offering clients and investors the ability to consider the corporate social performance from different standpoints. Furthermore, the notable differences in industry weights and ratings may be symptomatic of the agencies' dependence on companies and investors in a particular industry; thereby seeking to be flexible or 'adaptive' to client/investor demands. In terms of an explanation of the observed homogeneity, we argue there has been a noticeable rationalisation of rating agencies worldwide with a number of them merging, affiliating or being altogether acquired<sup>vi</sup>. Although this is a market-driven dynamic, less attention has been paid to what we would characterise as a *globalising* trend where firms such as CSR rating agencies also seek to expand and capture markets overseas, particularly in response to the 'mainstreaming' of ethical investment on a worldwide basis (Sandberg et al. 2009).

## **6. Conclusion and implications**

This paper sought to explore the methodologies used by a group of CSR rating agencies in evaluating the social performance of companies and in advising potential investors and clients. Prior studies did identify instances of similarities in rating methodologies whilst others had emphasised that significant differences appeared to persist. This has resulted in

debates about the need for greater standardisation, computational consistency and even the regulation of CSR rating practices. However, the empirical evidence was limited to specific CSR themes or only relied on publicly-available information. We therefore aimed to gather further exploratory evidence of practice from the agencies themselves. Our questions to the informants also focused on aspects not previously mentioned in the literature. Finally, we draw on Sandberg et al.'s (2009) conceptualisations relating to heterogeneity in the SRI movement as a tentative explanation of our findings; in particular, the levels at which such differences occur, and the reasons that may explain such differences.

Although we would have preferred to elaborate our discussion of these issues with more rating agencies, confidentiality concerns were often put forward to decline participation or to elaborate further. We nonetheless obtained sufficient insights on the way agencies have actually adopted exclusionary and positive criteria, the CSR themes of interest and the way weights are attached to the CSR themes within different industries. In effect, we find a combined picture of similarities and differences in the way the methodologies are structured and implemented on an industry basis, and also in terms of the way these methodologies are communicated. Unlike prior studies (e.g. Chatterji & Levine, 2008), we argue that a strategy of differentiation is not the only key force of interest. We also propose that there is a rationalisation within CSR ratings sector, which in turn leads to a more globalising influence in local settings and thereby fostering homogeneity. In this regard, we tentatively suggest that cultural or national factors may well appear to be on the wane.

Our paper's findings therefore add to the literature by suggesting that rating methodologies adopted by agencies display similarities as well as differences. From the perspective of sophisticated client-investors or market-related users of CSR ratings, the differences or

similarities may be less of a concern given their appreciation of the malleable nature of ratings. Yet, one might contend that the reliability and availability of the CSR ratings needs to be improved to attract more usage of SRI screens by the ethical investors (Dilla et al., 2016). We would also argue there remains an issue of transparency for other social, indirect or regulatory stakeholders. In particular, there has been increasing attention on, and acceptance of, the role of accounting, reporting and assurance practices in supporting the implementation of the sustainability agenda at the corporate, State and third sector level. This has become a critical issue in the aftermath of the 2016 Paris Agreement on Climate Change, the prominence of Sustainable Development Goals (SDGs) and the commitment to an enhanced transparency framework by a number of signatory countries. In such context, it becomes less tenable for CSR rating agencies to retain opacity and customisation if ratings are to be relied upon for transparency, monitoring and reporting purposes at national levels. Given the increased prominence of metrics in the CSR arena to compare social performance across firms and across countries, our findings highlight that the *'devil is very well in the detail'* in terms of how intermediary organisations, such as ratings agencies, operationalise CSR performance.

We acknowledge the following limitations for this study. A number of ratings agencies we contacted decline to participate in our study and amongst those who did respond, respondents were hesitant to provide more detailed insights or underlying motivations, often citing confidentiality as a means to ensure their competitive advantage. The relatively rapid change in the sector (mergers, acquisition and closure of agencies) also contributed to the challenge in obtaining a deeper access to some organisations. Therefore, we acknowledge that our findings, ensuing discussions and implications need to be considered in this light, and are in need of further validation. Our analysis therefore remains tentative but nonetheless of value in terms of suggesting avenues of analysis for further research. This issue of access has been

highlighted in prior studies (e.g. Sustainability, 2012) and a more in-depth access to different staff in agencies (similar to the in-depth study by Bessire and Onnee, 2010) would have certainly allowed us to explore the motivations underlying the use of specific methodologies. At this juncture, given the paucity of insights on the detailed aspects of ratings methodologies and the increasing importance of rating, it would be useful for future research to approach CSR rating agencies for a more in-depth understanding of the factors influencing the design and application of such ratings, potentially as part of an in-depth case study design. Potentially, Sandberg et al. (2009) does also offer a useful frame to explain why agencies might adopt different or similar methodologies and provide a starting point for a deeper analysis of the motivations and behaviour of CSR rating agencies.

## References

- Allouche, J. & Laroche, P. (2006), "Corporate Social Responsibility and Corporate Financial Performance: a Survey", in Allouche J. (eds), *Corporate Social Responsibility*, Palgrave Macmillan, London UK
- Archel, P., Husillos, J. & Spence, C. (2011), "The institutionalization of unaccountability: Loading the dice of corporate social responsibility discourse", *Accounting, Organizations and Society*, Vol. 36 No. 6, pp. 327-343.
- Ashforth, B. E., & Gibbs, B. W. (1990), "The double-edge of organizational legitimation.", *Organization Science*, Vol. 1 No. 2, pp.177-193.
- Bessire, B. & Onnee, S. (2010), "Assessing corporate social performance: Strategies of legitimation and conflicting ideologies", *Critical Perspectives on Accounting*, Vol. 21 No. 6, pp. 445-467.
- Carroll, A.B. (1979), "A Three-Dimensional Conceptual Model of Corporate Social Performance", *Academy of Management Review*, Vol. 4 No. 4, pp. 497-505.

- Carroll, A.B., & Shabana, K.M. (2010), "The business case for corporate social responsibility: A review of concepts, research and practice", *International Journal of Management Reviews*, Vol. 12 No. 1, pp. 85-105.
- Chatterji, A., Levine, D.I. & Toffel, M.W. (2009), "How Well Do Social Ratings Actually Measure Corporate Social Responsibility?", *Journal of Economics & Management Strategy*, Vol. 18 No.1, pp. 125-169
- Chatterji, A. & Levine D. (2008), "Imitate or Differentiate? Evaluating the validity of corporate social responsibility ratings", *Working paper, Center for responsible Business, UC Berkley*, Available at: <https://escholarship.org/uc/item/3sz7k7jc>
- Chatterji, A. & Levine D. (2006), "Breaking down the Wall of Codes: Evaluating Non-Financial Performance Measurement", *California Management Review*, Vol. 48 No. 2, pp. 29-51.
- Chelli, M., & Gendron, Y. (2013), "Sustainability ratings and the disciplinary power of the ideology of numbers", *Journal of Business Ethics*, Vol. 112 No. 2, pp. 187-203.
- Dahlsrud, A. (2008), "How Corporate Social Responsibility is defined: an Analysis of 37 Definitions", *Corporate Social Responsibility and Environmental Management*, Vol. 15 No. 1, pp. 1-13.
- D'Hollander, D., and Marx. A. (2014), "Strengthening private certification systems through public regulation: The case of sustainable public procurement", *Sustainability Accounting, Management and Policy Journal*, Vol. 5 No. 1, pp. 2-21.
- Dilla, W., Janvrin, D., Perkins, Jon. And Raschke, R. (2016) "Investor views, investment screen use, and socially responsible investment behavior", *Sustainability Accounting, Management and Policy Journal*, Vol. 7 No. 2, pp. 246-267

- Dillenburg, S., Green, T., & Erekson, H. (2003), "Approaching socially responsible investment with a comprehensive ratings scheme: Total social impact", *Journal of Business Ethics*, Vol. 43 No. 3, pp. 167–177.
- DiMaggio, P. J. & Powell, W. (1983), "The iron cage revisited" institutional isomorphism and collective rationality in organizational fields", *American Sociological Review*, Vol. 48, pp. 147-160.
- Entine, J. (2003), "The myth of social investing: A critique of its practice and consequences for corporate social performance research", *Organization & Environment*, Vol. 16 No. 3, pp. 352-368.
- Escrig-Olmedo, E., Munoz-Torres, M. J. & Fernandez-Izquierdo, M. A. (2010), "Socially responsible investing: sustainability indices, ESG rating and information provider agencies", *International Journal of Sustainable Economy*, Vol. 2 No. 4, pp. 442-461.
- Fassin, Y. (2009), "The Stakeholder Model Refined". *Journal of Business Ethics*, Vol 84, pp. 113-135.
- Gauthier, C. (2005), "Measuring corporate social and environmental performance: The extended life-cycle assessment", *Journal of Business Ethics*, Vol. 59, No. 1/2, pp. 199-206.
- Griffin, J. J. & Mahon, J. F. (1997), "The Corporate Social Performance and Corporate Financial Performance Debate", *Business and Society*, Vol. 36 No. 1, pp. 5-31.
- Hawken, P. (2004), *How the SRI industry has failed to respond to people who want to invest with conscience and what can be done to change it*, Natural Capital Institute: Sausalito
- Hedesström, M., Lundqvist, U. & Biel, A. (2011), "Investigating consistency of judgment across sustainability analyst organizations", *Sustainable Development*, Vol. 19 No. 2, pp. 119-134.

- Igalens, J. & Gond J-P. (2005), "Measuring Corporate Social Performance in France: A Critical and Empirical Analysis of ARESE data", *Journal of Business Ethics*, Vol. 56 No. 2, pp. 131-148
- Igalens, J., Dejean, F. & Akremi, A. (2008), "L'influence des systèmes économiques sur la notation sociétale", *Revue Française de Gestion*, Vol. 3, pp. 135-155.
- Knoepfel, I. (2001), "Dow Jones Sustainability Group Index: A Global Benchmark for Corporate Sustainability", *Corporate Environmental Strategy*, pp. 6-15.
- Knox, S. & Maklan, S. (2004), "Corporate social responsibility: Moving beyond investment towards measuring outcomes", *European Management Journal*, Vol. 22 No.5, pp. 508-516.
- Koehn, D. & Ueng, J. (2005), "Evaluating the Evaluators: Should Investors Trust Corporate Governance Metrics Ratings?", *Journal of Management and Governance*, Vol. 9, pp. 111-128.
- Malsch (2013), "Politicizing the expertise of the accounting industry in the realm of corporate social responsibility", *Accounting, Organizations and Society*, Vol. 38 (2), pp.149-168
- Margolis, J. D. & Walsh, J. P. (2003), "Misery loves companies: Rethinking social initiatives by business", *Administrative Science Quarterly*, Vol. 48 No. 2, pp. 268-305.
- Marquez, A. & Fombrun, C. J. (2005), "Measuring Corporate Social Responsibility", *Corporate Reputation Review*, Vol. 7 No. 4, pp. 304-308.
- Mattingly, J.E. & Berman, S.L. (2006), "Measurement of Corporate Social Action: Discovering Taxonomy in Kinder, Lydenberg, Domini Ratings Data", *Business and Society*, Vol. 45 No. 1, pp. 20-46.
- Mohr, L., Deborah, W. & Harris, K. (2001), "Do consumers expect companies to be socially responsible? The impact of corporate social responsibility on buying behavior", *The Journal of Consumer Affairs*, Vol. 35 No. 1, pp. 45-67.

- Miller, E., Buys L. & Summerville J. (2007), “Quantifying the Social Dimension of Triple Bottom Line: Development of a Framework and Indicators to Assess the Social Impact of Organizations”, *International Journal of Governance and Ethics*, Vol. 3 No. 3, pp. 223-237.
- O’Rourke, A. (2003). “The message and methods of ethical investment”, *Journal of Cleaner Production*, Vol. 11 No. 6, pp. 683-693.
- Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). “Corporate social and financial performance: A meta-analysis”, *Organization Studies*, Vol. 24 No. 3, pp. 403-441.
- ORSE (2012), *Guide des organismes d’analyse sociale et environnementale*, October. Available at:  
[http://www.orse.org/force\\_document.php?fichier=document\\_928.pdf&fichier\\_old=2012\\_version\\_finale.pdf](http://www.orse.org/force_document.php?fichier=document_928.pdf&fichier_old=2012_version_finale.pdf)
- Parguel, B., Benoît-Moreau, F., & Larceneux, F. (2011), “How sustainability ratings might deter “greenwashing”: A closer look at ethical corporate communication”, *Journal of Business Ethics*, Vol. 102 No. 1, pp. 15–28.
- Podnar, K. & Jancic, Z. (2006), “Towards a categorization of stakeholder groups: An empirical verification of a three-level model”, *Journal of Marketing Communication*, Vol. 12 No. 4, pp. 297-308.
- Rhodes, M. J. & Soobaroyen, T., (2010), “Information asymmetry and socially responsible investment”, *Journal of Business Ethics*, Vol. 95 No. 1, pp. 145-150.
- Sandberg, J., Juravle, C., Hedesstrom, T.M. & Hamilton, I. (2009), “The Heterogeneity of Socially Responsible Investment”, *Journal of Business Ethics*, Vol. 87, pp. 519-533.
- Scalet, S., & Kelly, T. F. (2010), “CSR rating agencies: What is their global impact?”, *Journal of Business Ethics*, Vol. 94 No. 1, pp. 69–88.



- Scott, M. E., Cocchi, D., & Campbell-Gemmell, J. (2014), “Defining a fit for purpose statistically reliable sustainability indicator.”, *Sustainability, Accounting, Management and Policy Journal*, Vol. 5 No. 3, pp.262-267.
- Semenova, N., & Hassel, L. G. (2014), “On the Validity of Environmental Performance Metrics”, *Journal of Business Ethics*, Vol. 132 No. 2, pp. 1-10.
- Sharfman, M. (1996), “The construct validity of the Kinder, Lydenberg, & Domini social performance ratings data.”, *Journal of Business Ethics*, Vol. 15 No. 3, pp. 287-296.
- Sustainability (2012), *Rate the Raters*, October. Available on: <http://www.sustainability.com/>
- Suchman, M.C. (1995), “Managing Legitimacy: Strategic and Institutional Approaches.” *Academy of Management Review*, Vol. 20 No. 3, pp. 571-610
- Wood, D. J. & Jones, R. E., (1995), “Stakeholder mismatching: a theoretical problem in empirical research in corporate social performance”, *International Journal of Organizational Analysis*, Vol. 3, pp. 229-267.

---

<sup>i</sup> [www.orse.org](http://www.orse.org)

<sup>ii</sup> [www.eiris.org/about-us/eiris-network/](http://www.eiris.org/about-us/eiris-network/)

<sup>iii</sup> Carroll (1979) proposed four dimensions of CSR namely economic, legal, ethical and philanthropic responsibilities. In a recent paper Carroll & Shabana (2010) put forward what they called the five most common dimensions of CSR: interest dimension, social dimension, economic dimension, the voluntary dimension and environmental dimension. Even if several definitions of CSR were otherwise developed (e.g. see the analysis of 37 definitions of CSR made by Dahlsrud, 2006), most of them agree on the threefold responsibility of the firm (Economic, Social and Environmental aspects) considering CSR as a company’s long term commitment to minimizing/eliminating any harmful effects on society and maximizing its long term beneficial impacts (Mohr et al., 2001).

<sup>iv</sup> Examples of general criteria are Corporate Governance and Risk & Crisis Management (Economic dimension), Environmental Policy & Management Systems (Environmental dimension) and Human Capital Development and Talent Attraction & Retention (Social dimension). With regard to industry-specific criteria, in the economic dimension, we can consider Anti-Crime Policy & Measures for the Banking sector and Market Opportunities for Electricity. In the environmental dimension, Biodiversity, Electricity Generation or Water-Related Risks are examples of industry-specific criteria for the electricity sector. Within the social dimension, Controversial Issues, Dilemmas in Lending & Financing are examples of industry-specific criteria for the banking sector while the strategy to Improve Access to Drugs or Products could be a specific criterion for Pharmaceuticals.

<sup>v</sup> We thank one of the anonymous referees for this pertinent point.

<sup>vi</sup> An illustration of this rationalisation is the number of agencies worldwide identified by Chelli and Gendron (2013), which a few years later declined significantly when we sought to confirm the number of active agencies. More recently Vigeo and Eiris, the French and the British CSR rating agencies, announced their merger into one single European agency.

## Appendix 1

	<b>Agency</b>	<b>Country</b>	<b>Founded</b>
1	BMJ Ratings	France	1993
2	Caer/Corporate monitor	Australia	2000
3	Champlain Research	France	2011
4	Covalence	Sweden	2011
5	E. Capital Partners	Italia	1997
6	Eco-Frontier Co.	South Korea	1995
7	Ecovadis	France	2007
8	Ecovalores	Mexico	2010
9	Eiris	UK	1983
10	Ethibel	Belgium	1991
11	Ethical Screening	UK	1998
12	Ethicx SRI Advisors	Sweden	1999
13	Ethifinance	France	2003
14	FundacionEcologia y Desarrollo (ECODES)	Spain	1992
15	GES Investment Services	Sweden	1992
16	Greeneye	Israel	NC
17	IMUG	Germany	1999
18	Inrate	Sweden	1990
19	KO-CSR	South Korea	2007
20	Management and Excellence (M&E)	Spain	2001
21	Oekom Research	Germany	1993
22	Reputex	Hong Kong	1999
23	SAM Research	Sweden	1995
24	Siris	Australia	2000
25	Solaron	India	2007
26	Sustainalytics	Netherlands	2008
27	The Good Bankers Co	Japan	1998
28	Vigeo	France	2002

---

## Appendix 2

### STUDY OF THE METHODS AND CRITERIA USED BY SUSTAINABILITY RATINGS AGENCIES (SRA) -LIST OF OUTLINE QUESTIONS

#### PART 1 EXCLUSION CRITERIA

Some approaches/methods of rating exclude certain industries for religious, moral or political reasons or for their involvement in activities that threaten the environment, the human health or the ecosystem.

- (a) Do you exclude some industries from your assessment process?
- (b) For which reasons do you exclude these industries?
- (c) Can you give some examples of the industries you exclude?
- (d) If a company is involved in one of the industries you exclude, what is the level or threshold of involvement (e.g. as a percentage of its total turnover) that once exceeded, the company is excluded?

#### PART 2 CSR DIMENSIONS

In order to assess the Corporate Social Responsibility (CSR) of the companies they rate, some Social Rating Agencies (SRA) identify and examine three or more dimensions (or themes) of CSR.

- (a) What are the dimensions/themes of the CSR that you consider in your assessment process?
- (b) Do these dimensions/themes have the same weight in your assessment regardless of the industry examined?
- (c) What is the weight (in percentage) of each dimension/theme of CSR in the final score of the firm assessed? All industries or Banks / Electricity / Pharmaceuticals

#### PART 3 SELECTION CRITERIA

In order to evaluate all the dimensions/themes of CSR, SRAs define, examine and check numerous criteria for each dimension. These criteria can be common to all the companies assessed regardless of the industry to which they belong (general criteria) or change according to the industry to which the company assessed belongs (industry-specific criteria)

- (a) Do you use general criteria? If yes, what is the weight of general criteria in each dimension/theme of CSR?
- (b) Do you use industry-specific criteria? If yes, what is the weight of the industry-specific criteria in each of these three industries? Banks / Electricity / Pharmaceuticals
- (c) Any other information or method which influences your selection of criteria.

**Table 1 Summary of Exclusion criteria (only for rating agencies using this approach)**

	Types of exclusion criteria	Examples of industries excluded	Threshold of involvement justifying exclusion
RA3	Ethical /religious and environmental (customized approach)	Weapons, nuclear, child labor	From 5% to 20% (most investors choose 5%)
RA4	Religious and ethical (combined with a customized approach)	Alcohol, Contraceptives, Tobacco, Gambling, GMO Food Production, Military army, Nuclear Power, Pornography.	Up to 5% but some industries are excluded no matter the level of involvement.
RA8	Ethical	Weapons, alcohols, Tobacco, etc.	Fixed by the client
RA9	Religious	Financial services, hotel industry.	From 5% to 100%
RA10	Religious, Ethical, Political, environmental and long term performance	Tobacco, weapons, nuclear power	From 5% to 20% depending on the client

**Table 2 Summary: weight (in percentage) of Social and Environmental themes by industry evaluated**

	Social			Environment		
	B	E	P	B	E	P
RA2	50	40	60	50	60	40
RA7	60	46	83	39	54	17
RA9	21	29	30	39	38	36
RA10	62	22	52	30	70	40

*B: Banking, E: Electricity and P: Pharmaceuticals*

**Table 3 Summary: weight of selection criteria by type, theme and industry**

	General criteria	Industry specific criteria
RA2	Social : 60-70 % Environmental : 40-60 %	Banks : Social 30 % – Environmental 50 % Electricity : Social 35 % – Environmental 60 % Pharmaceuticals : Social 40 % - Environmental 60 %
RA4	100%	100% (Products, Production process)
RA6	Roughly 2/3	Roughly 1/3 in each industry
RA7	40-50%	Banks: Economic 47 % - Environmental 70 % – Social 40 % Electricity: Economic 51 % - Environmental 83 % – Social 30 % Pharmaceutical: Economic 55 % - Environmental 50 % – Social 56 %
RA9	Environment : 40-50% Social : 30-50% CG : 100%	Banks: Environmental 52 % - Social 49 % Electricity: Environmental 58 % - Social 70 % Pharmaceutical: Environmental 54 % - Social 70 %
RA10	70-90%	10-30%

*For each theme (economic, social and environmental), evaluation is done based on both general and industry-specific criteria. Column 3 gives the weight of industry specific criteria in each theme within the three sectors, but the corresponding weight of general criteria is not given since our respondents only gave an average estimation of that weight for all the industries.*