Assessing companies' decent work practices: An analysis of ESG rating methodologies

Céline LOUCHE,* Guillaume DELAUTRE** and Gabriela BALVEDI PIMENTEL***

Abstract. *Environmental, social and governance (ESG) rating agencies, as non-financial data providers, have become central actors in the field of responsible investment. Although research has explored the construction of ESG metrics, little is known about how agencies evaluate decent work. Building on the analysis of six rating agencies, this article investigates how these actors measure and assess companies' performance in terms of decent work and related areas and identifies the challenges they face in this endeavour. The authors seek to better understand the capacity of responsible investment and ESG ratings to promote and improve decent work within companies.*

Keywords: *ESG rating, responsible investment, decent work, measurement, corporate social responsibility.*

1. Introduction

Decent work is increasingly considered to be an essential part of business responsibility. Promoted by the ILO since 1999, it is defined as productive work for women and men in conditions of freedom, equity, security and human dignity (ILO 1999). It aims to promote rights at work, encourage decent employment opportunities, enhance social protection and strengthen dialogue in handling work-related issues. Decent work is part of the Sustainable Development Goals (SDGs) established by the United Nations (UN) General Assembly in 2015 under SDG 8. In the last decade, the role of businesses in achieving decent work has also been reshaped by a growing number of initiatives, both regulatory (such

* Audencia Business School, email: clouche@audencia.com; ** ILO, email: delautre@ilo.org (corresponding author); *** Consultant, email: gabrielabpimentel@gmail.com. The research reported in this article received financial support from the French Government as part of its cooperation agreement with the ILO. The authors wish to thank Andrew Crane, Jake Reynolds and Fernando Messineo Libano for their helpful and constructive comments on earlier drafts of this article.

Responsibility for opinions expressed in signed articles rests solely with their authors, and publication does not constitute an endorsement by the ILO.

This article is also available in French, in *Revue internationale du Travail* 162 (1), and Spanish, in *Revista Internacional del Trabajo* 142 (1).

Copyright © The authors 2022
Journal compilation © International Labour Organization 2023

as the UK Modern Slavery Act 2015, the French “Duty of Care” Act of 2017 and EU Directive 2014/95/EU) and voluntary (such as the Global Reporting Initiative or the UN Global Compact). All these initiatives require companies to disclose publicly non-financial information, including on decent work.

In this context, it is central to evaluate companies’ commitment to, and practice of, decent work. The question has become even more relevant given the growth of responsible investment (RI). RI can be defined as “an approach to investing that aims to incorporate environmental, social and governance (ESG) factors into investment decisions, to better manage risk and generate sustainable, long-term returns” (PRI 2018). ESG-related data, including on decent work, are essentially provided by ESG rating agencies. In the last 20 years, these agencies have caught the attention of scholars, who have studied their role as metric providers (Chatterji, Levine and Toffel 2009), their methodologies (Eccles, Lee and Stroehle 2020; Eccles and Stroehle 2018; Penalva-Icher 2016; Stroehle 2018) and the measurement challenges they face (Berg, Kölbel and Rigobon 2022; Chatterji et al. 2016). Although these studies have provided rich insights, little is known about the specificities of each ESG dimension, especially as regards decent work.

Accordingly, this article explores two main questions: (1) How do ESG rating agencies assess companies’ performance on decent work?; and (2) What are the challenges they face in this assessment? Our research thus aims to develop a better appreciation of RI and agencies’ capacities to promote decent work.

To this end, we conducted an exploratory study among six major ESG rating agencies. The study is based on 18 interviews with agency representatives, and ESG and decent work experts, and on multiple sources of secondary data. Our results show that, although agencies provide a broad coverage of the subject of decent work, including all the fundamental rights at work as defined by the ILO, they face significant challenges in assessing decent work as a stand-alone concept. We identified three categories of challenges to this assessment: first, the nature of ESG data, notably the difficulty of gathering reliable information and properly quantifying it; second, the nature of decent work issues, which are perceived as subjective, sensitive and context-dependent; and third, the nature of the ESG rating agencies, which are bound to the principles of materiality and take part in a highly competitive and changing environment. Our study reveals that, although agencies have a role in promoting decent work, especially regarding disclosing information, they remain relatively constrained by technical, ethical and structural factors.

The remainder of this article is organized as follows. The second section examines the development of the concept of RI and the ESG rating agency market, outlining the factors that influence agencies’ measurement and evaluation processes. The third section sets out our research methods, while the fourth section presents our results. The fifth section discusses our findings and concludes by pointing to future areas for possible research.

2. Responsible investment and ESG rating agencies

2.1. Responsible investment

Once a niche subject, RI has become part of mainstream financial markets (Dumas and Louche 2016). The *Global Sustainable Investment Review 2020* estimated the RI market at over US$35 trillion (Global Sustainable Investment Alliance 2021). The UN-supported Principles for Responsible Investment (PRI) had 3,826 signatories in 2021, including the largest institutional investors and asset managers worldwide, representing over US$121 trillion under management.[1]

Although RI takes many forms and is subject to many approaches, four main strategies have been identified in the literature: avoidance (the use of exclusionary criteria to avoid businesses or practices regarded as unacceptable or harmful to society, such as alcohol, tobacco industries and involvement in human rights violations); inclusion (the use of thematic screening on areas such as climate change, water or access to medicine); relative selection (the use of best-in-class approaches to select the best ESG-performing companies); and engagement or shareholder activism (making use of one's ownership position to actively influence the company through proxy voting, shareholder resolutions or dialogue, among others) (Sjöström 2008). All these strategies rely on access to ESG data.

2.2. ESG rating agencies

ESG rating agencies assess and compare companies on the basis of their corporate social responsibility (CSR) performance by collecting, aggregating and interpreting large quantities of information. They have become prominent actors in the RI field and now play a central role in constructing and operationalizing ESG data (Zarlowski 2007). Their clients are mainly private and institutional investors, and asset managers.

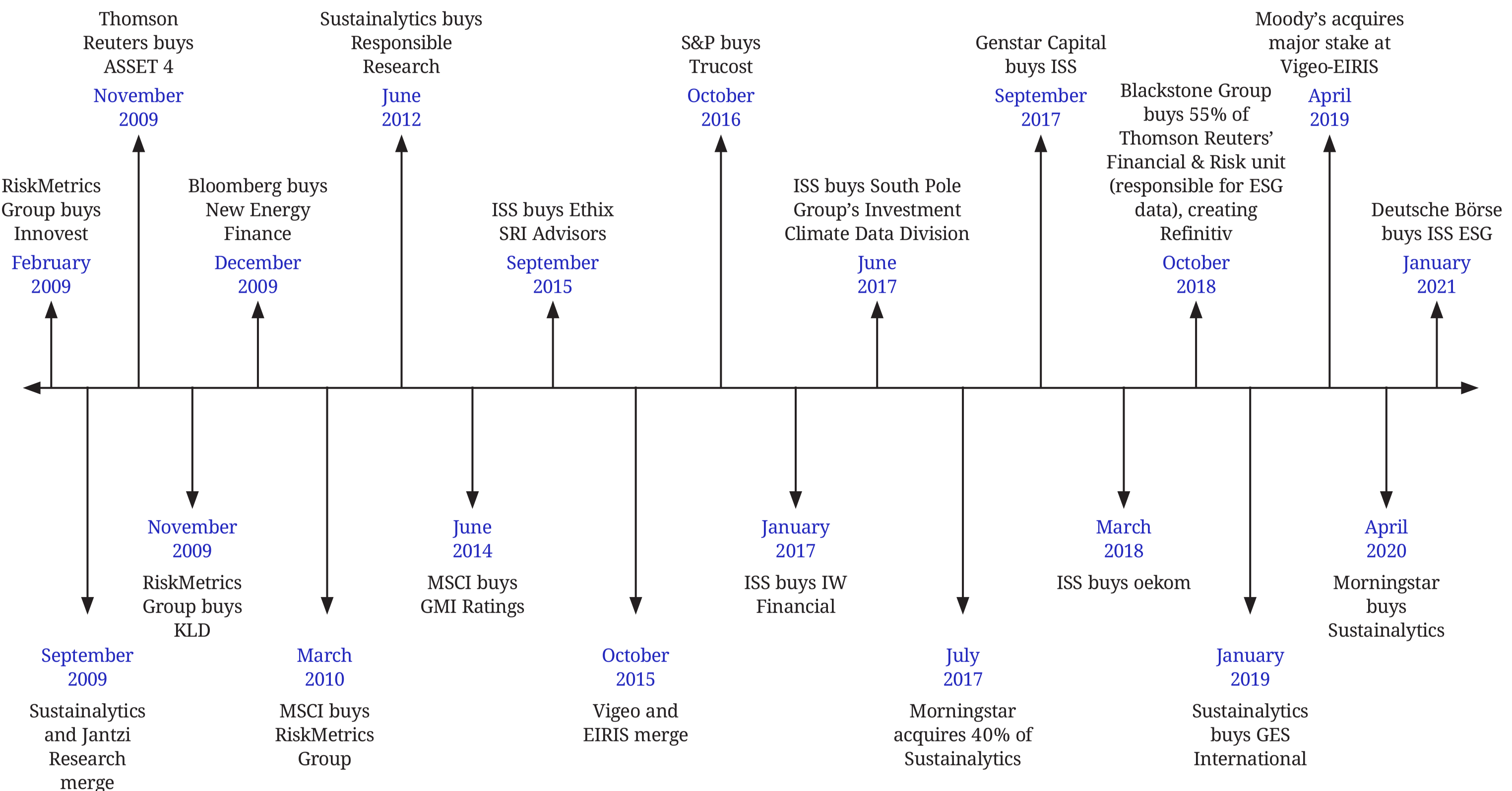
Many of these agencies were created between the 1980s and 1990s in Europe and North America. In the last 30 years, the market has undergone strong consolidation, which seeks to increase agencies' capacity to expand the universe of rated companies and reach financial stability (Dimmelmeier 2020; Avetisyan and Hockerts 2017; Brown and Wallace 2018) (see figure 1). Most of the rating agencies have been acquired by US companies such as Morgan Stanley Capital International (MSCI), Morningstar or Moody's Investor Services, leading to a convergence of mainstream financial actors and ESG actors (Novethic 2018; Nauman 2019).

2.3. ESG rating agencies as metrics providers

Agencies have designed in-house methodologies to assess companies' ESG performance. Several studies have investigated these methodologies and the metrics elaborated by the agencies. Empirical work has focused mainly on the aggregated ESG score (Chatterji et al. 2016; Chatterji, Levine and Toffel 2009; Widyawati

[1] See https://www.unpri.org/about-us/about-the-pri, accessed on 7 May 2022.

Figure 1. Consolidation of the ESG rating agencies' market



Source: Adapted from Brown and Wallace (2018, 6).

2021) and the environmental dimension of ESG assessments (Chatterji, Levine and Toffel 2009; Semenova and Hassel 2015), largely leaving out the social dimension in general, and labour-related subdimensions in particular.

Although all rating agencies assess how well companies manage ESG risks and opportunities, their approaches differ considerably. The existing literature points to three main factors that influence agencies' measurement and evaluation processes, namely: the agency's social origins, its understanding of materiality and its methodological choices.

The social origins of ESG rating agencies refer to the history of rating organizations. This history infuses an agency's understanding and conceptualization of ESG indicators and, consequently, how it is measured (Eccles, Lee and Stroehle 2020). The approach taken to evaluate companies, including the type of indicators they use and how they construct them, is directly influenced by how an agency defines its own mission. Eccles and Stroehle (2018) identify two types of agencies. The first is defined as "value-driven" and focuses on ESG information that is financially relevant. In this case, the ESG rating reflects the capacity to create corporate value and shareholder return. The second type of agency is defined as "values-based", focusing on sustainability and societal impacts. Consequently, value-driven agencies tend to rely more on quantitative and performance-based metrics, while values-based agencies tend to favour a qualitative and policy-related approach (Eccles and Stroehle 2018). The social origins of rating agencies help explain idiosyncrasies in the methodological decisions taken by data providers (Stroehle 2018). Nevertheless, the consolidation of the industry suggests a growing trend towards a more value-driven approach (Bouten et al. 2017; Eccles, Lee and Stroehle 2020).

The concept of "materiality" originates from the field of financial analysis. Information is considered material if there is "a substantial likelihood that the disclosure of the omitted fact would have been viewed by the reasonable investor as having significantly altered the 'total mix' of information made available" (citing the Supreme Court of the United States, Eccles and Stroehle 2018, 9). The same concept is used in RI to assess which factors of ESG evaluations matter the most for investors (Eccles, Lee and Stroehle 2020). Materiality is a highly controversial concept, since it implies the prioritization of information seen as relevant for investment decisions, working as a filter through which management values information (Hoogervorst 2017). For most actors in the finance industry, material ESG issues are those that have a significant impact on revenue and return on capital and, therefore, have the greatest influence on a firm's ability to generate shareholder value (Eccles and Serafeim 2013).

Several studies have highlighted problems of divergence between different ESG ratings (Chatterji et al. 2016; Christensen, Serafeim and Sikochi 2022; Delmas, Etzion and Nairn-Birch 2013; Berg, Kölbel and Rigobon 2022) and of low levels of correlation across ESG ratings (Chatterji, Levine and Toffel 2009; Kölbel et al. 2020). Chatterji et al. (2016) showed that commensurability – that is, "how raters measure the same constructs" (p. 1600) – among rating agencies is low, even when adjusted for explicit differences in their theorization of ESG – that is, "the beliefs raters have about what being socially responsible means" (p. 1599).

They concluded that most assessments provided by agencies have a high potential for measurement error. In line with previous findings, Berg, Kölbel and Rigobon (2022) identify three primary sources of divergence among ESG ratings: "scope divergence", referring to the variation in the range of issues considered; "weight divergence", as agencies give different importance to the same issues; and, more importantly, "measurement divergence", referring to the differences found in the indicators used. At a more disaggregated level, studies show that the correlation of scores related to social dimensions (including labour and rights at work issues) is generally lower across rating agencies than the correlation of scores on environmental dimensions. According to these studies, the correlation can be negligible – and even negative – depending on the pair of agencies under study (Berg, Kölbel and Rigobon 2022; Dorfleitner, Halbritter and Nguyen 2015). Such differences and inconsistencies are likely to lead to significantly diverse investment recommendations, create confusion and even misinform investors (Delmas, Etzion and Nairn-Birch 2013). That is why several business actors and academics have called for a move towards more regulation and standardization in ESG measurement (Berg, Kölbel and Rigobon 2022; MEDEF, Afep, Cliff and C3D 2019).

3. Research method

We chose a qualitative research design with a view to developing an in-depth understanding of the assessment of decent work for the purposes of RI. Qualitative methods are particularly effective in gathering rich information on phenomena that have been previously ignored in the literature (Creswell 1998).

3.1. Data collection

Our study is based on the analysis of six major ESG rating agencies (see table 1). Primary and secondary data were collected from various sources and we conducted 18 semi-structured interviews, 10 with agency representatives and 8 with experts working at or in collaboration with the ILO, universities or ESG-specialized organizations (see table 2).[2] Interviews included questions on (1) the interviewee's organization and their background, (2) their understanding of decent work, (3) their organization's methods of assessment and the challenges faced in evaluating labour-related indicators and (4) their thoughts about RI and decent work. We also collected secondary data from the agencies, including methodology documents, press releases, reports and sample company profiles. Additionally, we consulted publicly available information from other organizations such as the UN Global Compact, stock markets, Novethic[3] and PRI, as well as media sources.

[2] For confidentiality reasons, each individual interview was assigned a code. IntA indicates interviews with academics, IntE with experts and IntRA with rating agencies.

[3] A branch of the French Deposits and Consignment Fund (*Caisse des dépots et consignations*), acting as a media outlet to provide information and training on, and encourage action towards, sustainable transformation among financial actors.

Table 1. ESG rating agencies

ESG rating organization	Number of companies (worldwide)	Number of ESG analysts (approx.)	Number of ESG indicators	Headquarters	Indices
Sustainalytics (Morningstar)	11 000	250 (out of 650 employees)	220 indicators and 450 fields	United States	Global Sustainability Signatories Index; Jantzi Social Index; also subsidizes the production of the two FTSE Russell ESG Indices: FTSE Developed ESG Index and FTSE Emerging ESG Index
VigeoEiris (Moody's)	8 000	140 (out of 245 employees)	330 indicators	United States	8 Euronext indices and Ethibel Sustainability Index
MSCI	8 500	185 research analysts	56 indicators, 230 datapoints	United States	ESG rating covers companies included in a total of 22 MSCI indices
Refinitiv	9 000	150 research analysts	450 data points, 70 analytics	United States	–
RobecoSAM (S&P Global)	4 700	No information	An average of 100 data points, 23 criteria per company	United States	Used to inform the Dow Jones Sustainability Index
ISS ESG	8 000 (aiming to achieve 10 000 by the end of 2020)	200 research analysts (out of 2 000 employees)	An average of 100 criteria per company out of a pool of 800+ indicators	Germany	Supports the following indices: Solactive ISS ESG Screened Index Series; Solactive ISS Low Carbon; STOXX Low Carbon Index Series; OMX Stockholm 30 ESG Responsible Index; Global Challenges Index; SIX Sweden ESG Selection Index

Source: Authors' compilation based on publicly available information from the agencies' websites and documents that were accessed between May and September 2020.

Table 2. Interview overview

Category	Organization	Function	Duration (minutes)	Interviewee codes
Decent work experts	ILO	Social finance/enterprises	71	IntE1 to IntE4
		Enterprises/CSR team	60	
Decent work and ESG experts	ILO, ICFTU, Ethibel	Member of the Ethibel Register Committee	69	
ESG expert	United Nations' Principles for Responsible Investment	Manager, decent work programme	90	
Academics experts on ESG	University	Professor in Ethics and Economics	84	IntA1 to IntA4
		Research fellow	64	
		Professor in Applied Economics	70	
	Think Tank/University	Executive director	68	
Rating agencies	ISS–ESG	Senior associate, ESG ratings	84	IntRA1 to IntRA10
		Analyst	65	
	MSCI	Analyst	19	
	Refinitiv	Director	42	
		Head ESG	49	
	RobecoSAM	Managing director	83	
	Sustainalytics	Executive director	41	
	VigeoEiris	Analyst	87	
		Director	115	
		Analyst	58	

Source: Authors' compilation.

3.2. Data analysis

Prior to the interviews, we organized our secondary data in such a way as to gain an in-depth understanding of the selected agencies and their frameworks relating to decent work. We compiled the agencies' decent work criteria, assessment methodologies and sources of information in tables, to which we later added the information gathered in the interviews. This first step helped us understand how decent work was assessed and allowed us to compare the agencies.

The second step consisted in analysing the interviews that we carried out. Using NVivo, we applied a comprehensive coding procedure (Creswell 2013; Corbin and Strauss 2008). In line with our inductive approach, and to ensure an open-minded analysis, the initial data coding was based on open codes. We then moved back and forth between data and emerging concepts to reach a higher level of data coding. This involved axial coding, where we compared first-order codes with one another, looking for patterns and themes to create second- and third-order constructs (Gioia, Corley and Hamilton 2013). At this stage, we also used a set of a priori codes based on the literature, especially regarding the challenges linked to ESG data. The axial coding was carried out by one researcher and put to test by the other researchers in a series of meetings. Throughout our analysis, we triangulated interview material with secondary data, to ensure the robustness of our coding (Golafshani 2003).

4. Results

Measuring decent work is an essential part of ESG ratings. All agencies, without exception, assessed companies' policies and practices in this domain. Most of them considered the ILO's four fundamental principles and rights at work – freedom of association and right to collective bargaining, non-discrimination, abolition of child labour and elimination of forced labour – and other key issues such as health and safety, and social compliance in supply chains. We estimate that indicators related to decent work represent approximately 20 to 35 per cent of the overall ESG score. Although this is just an approximation, it indicates the importance of decent work in ESG ratings. However, as we will see later, our analysis reveals the difficulties encountered when assessing and trying to understand companies' global approach to decent work.

In this section, we first analyse the way in which ESG rating agencies approach decent work, highlighting the factors that generate variations in the assessment. We then investigate seven specific indicators of decent work and highlight how agencies approach them differently. Lastly, we explore the challenges faced by agencies in assessing decent work.

4.1. Variety of approaches to assess decent work

4.1.1. Diffuse understanding

The agencies that we analysed integrated dimensions of decent work in their assessments but did not refer to the concept itself. For most of the interviewees, the concept remained abstract: "I often have the feeling that it is something very

abstract. We can discuss about it, you can have international conventions on this, but what does it mean concretely?" (IntRA6).

The agencies preferred to use labels which, according to them, are better understood by investors, such as "human rights", or more business-oriented terms, such as "human capital management", "human resources" or "employee engagement". One interviewee noted that "decent work sounds a bit 'activisty'" (IntE4). However, they all referred to ILO standards and the UN Global Compact. One academic expert (IntA1) highlighted the importance of the "opposability" of the criteria to international norms and instruments, such as the Universal Declaration of Human Rights and the ILO's fundamental Conventions.[4]

4.1.2. Fragmented approach

It is difficult to have a clear understanding of how ESG rating agencies measure decent work. Due to business confidentiality, the information available is limited. More importantly, the measurement of decent work is fragmented and not always visible. We found decent work issues in categories as diverse as those of human rights, human resources, business relationships or supply chains. Even then, those issues were not always explicitly mentioned or they were sometimes hidden in subcategories. In addition, labels used to refer to the same dimensions of decent work differed from one agency to another.

4.1.3. Modular evaluation

Within each agency, the number of indicators used to assess decent work varied substantially, depending not only on the company assessed but also on the different aspects evaluated in that company. The number of criteria used also varied by sector and region. Indicators could be activated or deactivated based on perceived risks and relevance. IntRA7 gave the following example: "We tailor the choice of criteria … in the chemical sector, we will have something more about hazardous waste and so on, that we won't have in banks or in other [organizations]. We use criteria that are industry-specific."

Similarly, an indicator can be activated when assessing one part of a company's business and deactivated when assessing another. For example, IntRA3 informed us that, when a company originated from a region where legislation against child labour was regarded as duly enforced, child and forced labour were not included for direct employees of the company but only in the supply chain section.

4.1.4. Different weighting systems

Agencies apply a weighting system to their indicators based on in-house rules, which can vary by sector, location or other characteristics. For instance, IntRA7

[4] For more information, see https://www.ilo.org/global/standards/introduction-to-international-labour-standards/conventions-and-recommendations/lang--en/index.htm. As of 2022, the ILO identifies 11 fundamental Conventions. However, please note that at the time of writing, occupational safety and health (OSH) was not yet included in the list of Fundamental Principles and Rights at Work. ILO constituents adopted a resolution to include OSH in the list at the 110th Session of the International Labour Conference in June 2022.

explained that “social issues are more highly weighted in industries where it is more service-oriented”.

Another interviewee mentioned the example of human resource indicators: “For software and IT, we assess the respect and management of working hours, but we do not do that for broadcasting companies because they are not really subject to this issue.” (IntRA9). The weight can also depend on the availability and quality of data: “We don’t want to give a high weight for something that could then be a little bit wrong” (IntRA7). However, rating agencies did not always clarify the details of their weighting systems.

The underlying approach to materiality, which is directly influenced by the social origins of rating agencies, also influences their weighting system. For example, one agency indicated that the weight of each indicator at the industry level was based on its past impact on financial performance over time (measured by the level of correlation between the two measurements). Two other agencies included in their weighting system the risk exposure related to companies’ business models, products, geographical position, size and reliance on public contracts or outsourced production, among others. These agencies were more value-driven and, accordingly, increased the weight of indicators with a direct impact on the financial value of a firm. In contrast, one agency integrated norms-based criteria in its weighting system, revealing a more values-based orientation. Criteria relating to the nature of risk (for example, considering whether risks are connected to universally agreed human rights) or the specific vulnerabilities of a company’s stakeholders (such as operations in particularly risky environments for workers’ rights and working conditions) were examined alongside criteria with a greater bearing on financial outcomes, such as reputation, human capital, operational efficiency and legal security.

4.1.5. Varying evaluation of controversies

An ESG controversy is created by information of a public nature that places a company in the media spotlight and holds it liable for non-compliance on one or several rating indicators. Controversies usually refer to past or ongoing scandals, such as cases of toxic waste spill, human rights violations and corruption. ESG rating agencies identify controversies by continuously researching multiple media sources and publications by NGOs.

While all the agencies in our study included controversies in their assessments, their scoring methods varied significantly, since they built on several layers of interpretation. Controversies were assessed based on their level of severity, which depended primarily on the definition and understanding of the incident itself (“Is this something local or global? Does it involve top management?” (IntRA7); “Is there [a] death?” (IntRA9)). The scale of the impact on the company also played a role in assessing the severity of a controversy: “Was there a condemnation, a legal proceeding or is it just an allegation?” (IntRA9). In addition, rating agencies considered the frequency of controversies at the company level – that is, whether it was a one-time event or a repeated occurrence. Lastly, they evaluated other dimensions, such as the responsibility of the company (namely, for monitoring failures, breach of company policies or legislation and court decisions, among

others), financially material criteria (such as fines, reputation and production stoppages) and the quality of the company's response to the controversy. It emerged that the normative criteria were relatively marginal in this regard and were only reviewed by one values-based agency. In this case, the criticality (was it related to a universally recognized right?), the scale (how many stakeholders were affected?) and the irreversibility of the controversies were considered.

4.2. Analysis of decent work issues

In this section, we analyse the decent work issues covered by rating agencies. We grouped these issues into seven categories. The first three categories refer to the ILO fundamental principles and rights at work, grouped into freedom of association and the effective recognition of the right to collective bargaining, the elimination of discrimination in respect of employment and occupation, and the elimination of child labour and forced labour. Another three categories refer to other ILO standards relating to occupational safety and health, working conditions, and human resources and employment management. The last category focuses on social compliance in the management of supply chains. Table 3 provides insights into these seven categories, including examples of indicators used by rating agencies to assess decent work.

Freedom of association and the right to collective bargaining were mentioned by all agencies, but their interpretation and measurement varied substantially. As one interviewee noted, this indicator's objective was to "look at the extent to which the enterprise respects trade union freedom [and] collective bargaining rights and promotes collective bargaining right[s] … but then, what are the separate criteria?" (IntA1). It remained unclear which indicators most agencies used and what their scope was – that is, whether they applied to all rated companies or only to some. A minority of agencies deemed this indicator to be part of the dimension dedicated to human rights, while most identified it with "labour management" or "human capital". Those indicators were generally measured on the basis of policy commitments, implemented actions, and monitoring and grievance mechanisms. Some agencies also gave positive ratings to companies if they participated in the UN Global Compact or had formal bilateral commitments through international framework agreements (IFAs). Box 1 provides an illustration of how policy commitments were evaluated by one of the agencies studied. Some agencies used quantitative indicators, such as the number of employees covered by independent trade unions or the coverage of collective bargaining agreements. Interestingly, two agencies claimed to have a specific focus on companies operating in countries with severe restrictions to workers' collective rights. Thus, companies were asked about possible alternative worker participation measures that could be implemented in such countries to circumvent this structural problem. A minority of agencies chose a different approach based on a limited number of indicators, such as the number of labour disputes, work stoppages and strikes reported by the company. In these specific cases, it was suggested that industrial relations were seen as a risk factor for the business, leading to higher costs and lower profitability in the short term.

Non-discrimination and, more generally, the promotion of diversity in companies were widely examined by rating agencies. However, they differed

Table 3. Examples of decent work indicators used by ESG rating agencies

Groups	Categories	Coverage	Scope	Evaluation mode	Examples of indicators
ILO Fundamental Rights	Freedom of association and collective bargaining	All agencies	Respect of the freedom of association and collective bargaining	Mainly qualitative, based on policy commitments, with some quantitative indicators	Guarantee the freedom of association and collective bargaining Collective bargaining agreements Judiciary processes and campaigns Risk management and risk exposure to work stoppages and strikes Policy to promote labour relations or encourage employee participation Trade union representation
	Non-discrimination	All agencies	Mainly focused on diversity; limited consideration of other forms of discrimination	Mainly qualitative, based on policy commitments, with some quantitative indicators	Non-discrimination policy Diversity and inclusion programmes, commitments and data Diversity and gender equality agreements signed with trade unions or public authorities Policy to ensure equal remuneration and/or equal opportunities in recruitment
	Child labour and forced labour	All agencies (not always explicit)	Elimination of child labour and forced labour	Qualitative, based on policy commitments	Policy on child labour and forced labour Human rights policies and commitments Human rights due diligence process Human rights disclosure
Other ILO standards	Occupational safety and health	All agencies	Commitments, implementation strategies and certifications	Quantitative and qualitative	Health and safety policy Improvement of health and safety conditions Total recordable injury frequency rate Process safety events Work-related fatalities for employees and contractors Occupational illness frequency rate Judicial proceedings Policy to ensure work–life balance

(continued overleaf)

Table 3. Examples of decent work indicators used by ESG rating agencies *(concl.)*

Groups	Categories	Coverage	Scope	Evaluation mode	Examples of indicators
	Working conditions	All agencies except two, but not explicitly mentioned	Commitment to comply with applicable wage and hours legislation	Mainly qualitative, based on policy commitments, with some quantitative indicators	Working conditions policy Quality of remuneration systems Respect and management of working hours Employment security and types of employment
	Human resources and employment management	All agencies	Commitment and implementation of career management, but also – for a few – restructuring management processes	Quantitative and qualitative	Policy for human capital development Career management, promotion of employability and development programmes Training and education Human capital return on investment Return on employee development investment Internal mobility Individual performance appraisal Long-term incentives Talent attraction and retention policy Employee turnover rate Percentage of temporary workers Responsible management of restructurings Judicial proceedings
	Supply chains	All agencies	Scope, means and coverage of suppliers' policies	Mainly qualitative, based on policy commitments, with some quantitative indicators	Scope and quality of social supplier standards Conflict minerals policy Integration of ESG in supply chain management and strategy Promotion of social and economic development Suppliers' code of conduct, certifications and standards Risk exposure (identification of critical suppliers) and risk measurement Signatory to sector-specific frameworks

Source: Authors' compilation.

Box 1. Evaluation of a company's commitment to freedom of association (example)

For this illustration, we draw directly from the data of one of the rating agencies analysed. The agency measured companies' policy commitment to freedom of association on three levels: the degree of disclosure, the coverage of all duties related to a given issue and the internal support granted to the commitment inside the company.

Company A was a European company in the automotive industry. The agency had given the company the maximum score on policy commitment because: (1) the company had signed IFAs and other agreements with a major global union federation, specific agreements with unions in some countries of operation and the UN Global Compact; (2) the company's commitment under the IFA was detailed, addressed most of its responsibilities and explicitly supported the implementation of the ILO Freedom of Association and Protection of the Right to Organise Convention, 1948 (No. 87), Right to Organise and Collective Bargaining Convention, 1949 (No. 98) and Workers' Representatives Convention, 1971 (No. 135); and (3) the company's commitment applies to all levels of the firm and is supported by senior management, and the company collaborates with union representatives.

A company would receive a low score in this area if it did not demonstrate any public commitment or take part in any initiative.

significantly in the extent of their non-discrimination assessment and on the grounds for discrimination they considered. In accordance with ILO Discrimination (Employment and Occupation) Convention, 1958 (No. 111), half of the agencies explored multiple grounds for discrimination, such as gender, ethnic origin and age. The evaluations were mostly based on the formal commitments (such as discrimination policies) made by the companies, and the existence of concrete actions to fight discrimination through training, recruitment or other management systems. However, not all grounds for discrimination were treated equally. Indeed, gender equality and, to a lesser extent, disability were often given more importance than ethnicity or age. For some agencies, gender was the sole form of discrimination evaluated. In general, though, the assessment of gender discrimination was more detailed, including quantitative indicators on the number of women employees, access to management, executive positions and pay gaps.

Child labour and forced labour were less explicitly covered by agencies than the other ILO fundamental rights and principles. Only two agencies mentioned these indicators in their analytical grids, mostly through the evaluation of public engagement and instruments (for example, codes of conduct). In addition, these indicators seemed to be activated only for companies with operations in countries or industries considered to be at risk. However, the topic was usually raised in the evaluation of supply chains (see below) and was also monitored through screening for controversies.

Safety and health criteria were analysed by all agencies and applied to most sectors. However, the type and number of indicators used to assess them varied by sector. Three of the agencies prioritized policy commitments, implementation

strategies and certifications, such as OHSAS 18001 and ISO 45001. Another assessed safety and health by the number of judiciary procedures relating to this topic. Some also reported using indicators on occupational injuries, illness and fatalities. Interestingly, only one agency looked for information linked to the prevention of stress-related diseases and mental health.

Albeit to a lesser extent, assessment also included other criteria on working conditions. Two agencies explicitly mentioned assessing wages and overtime in their methodology documents. One evaluated the quality of remuneration systems, including their transparency and objectivity, as well as the management of working hours (for example, limits on working long hours and compensation rules for atypical schedules or overtime). Another agency evaluated the companies' compliance with applicable wage and working-hour legislation, and their adoption of industry-specific codes of conduct. One agency was starting to evaluate fair wages, although it had not yet integrated this indicator in its scoring system because of poor data availability. Two agencies considered the issue of working hours through the lens of access to flexible working time arrangements and work–life balance. It was, however, difficult to measure the extent to which these criteria were universally applied.

Among the myriad of disaggregated social indicators assessed by rating agencies, we identified a group of indicators related to human resources and employment management. These indicators related more specifically to career management, such as training systems, the existence of (individual) performance appraisal and talent identification mechanisms. Only half of the agencies specifically emphasized employment security, and this was mostly through indicators related to job turnover or the use of temporary employment. Two agencies evaluated the management of restructuring processes by analysing elements like the anticipation of reorganization, the involvement of staff representatives or the assistance provided to affected employees.

Social compliance in supply chain management was widely recognized as an important topic, but its treatment varied significantly. Four agencies provided reasonable details on the way they evaluated decent work in supply chains, which included the scope and coverage of their suppliers' policies. Agencies considered a wide range of issues, such as the ILO's four fundamental principles and rights at work, safety and health, (living) wages, working time, violence and harassment. In many cases, agencies considered a greater number of issues when evaluating suppliers' staff policy than they did for the rated companies' own policy. The agencies also aimed to evaluate the quality of existing procedures, such as social audit systems, procedures in the event of non-compliance, and internal training for procurement managers. Some agencies conducted special inquiries into specific industry-related topics, such as conflict mineral policies and programmes.[5] However, none of the agencies mentioned which supply chain levels (tiers) were taken into consideration.

---

[5] "Conflict minerals" can be defined as minerals sourced from politically unstable areas which are traded to finance armed groups, fuel forced labour and other human rights abuses, and support corruption and money laundering. They include minerals such as tin, tungsten, tantalum and gold, also referred to as 3TG, which can be used in everyday products, such as mobile phones and cars, or in jewellery. Conflict Minerals Regulation: The regulation explained (europa.eu)

Lastly, two issues related to decent work received very little attention. The first was the economic and social impact of companies' investments and activities on local communities and businesses. The responsibility of multinational companies in this regard has been recognized in the ILO Tripartite Declaration of Principles concerning Multinational Enterprises and Social Policy (see Articles 17 to 20).[6] However, only one agency included this criterion in its methodology. The second issue to receive little attention was social protection – one of the four aims of decent work – which was rarely mentioned in the agencies' methodologies. Access to social protection is a prerogative of governments and public authorities in most countries, which means that the scope, coverage and quality of social protection systems are highly uneven across the globe (ILO 2017). A few pioneer multinational enterprises have decided to provide innovative social protection benefits and services across the countries in which they operate, such as death, accident and disability insurance, paid maternity and paternity leave, healthcare and meal subsidies to their employees (for concrete examples, see Bourguignon and Mias 2017; Sekerler Richiardi and Arbo 2019).

4.3. Challenges of measuring decent work

This section focuses on the structural, technical and ethical challenges of measuring companies' performance in terms of decent work. We organized these challenges into three categories concerning the nature of ESG data, the nature of decent work issues and the nature of ESG rating agencies.

4.3.1. The nature of ESG data

This refers to the type of data required and the methodologies used to assess decent work. We identify three closely interrelated challenges in this regard. The most important of these is the quality of the data in terms of their reliability, comparability and availability. Indeed, ESG rating agencies rely heavily on publicly available information provided by the companies they rate. Research has shown that reporting practices differ greatly among companies, both in terms of the quantity and quality of the information (depending on factors such as company size, profitability and sector) and of the influence of institutional country-related factors, including access to freedom of expression and to media (Ali, Frynas and Mahmood 2017; Marquis, Toffel and Zhou 2016; Fifka 2013; Lucchini and Moisello 2017; Fortanier, Kolk and Pinkse 2011). Although the quality of data concerned all the dimensions of CSR, the interviewees noted that it particularly affected decent work: "This is one of the biggest challenges for the rating agencies, to get proper data" (IntRA4); "It is seldom that you would find very clear [evaluations] on the quality and the effective respect of labour rights" (IntRA6). Although agencies adopted different practices, companies were usually downgraded in the absence of information: "The lowest grade means that there is no information" (IntRA3).

[6] ILO, Tripartite Declaration of Principles concerning Multinational Enterprises and Social Policy, adopted at the 204th Session of the Governing Body (Geneva, November 1977) and amended at its 279th (November 2000), 295th (March 2006) and 329th (March 2017) Sessions.

When data were available, the question of reliability and exhaustiveness was raised. Many analysts mentioned the difficulty of interpreting the information provided: "Companies are not always fully transparent and fully honest in their communications" (IntRA10); "You are overwhelmed with very positive company information, that they promote of course themselves as being very responsible" (IntRA6). Previous research has indicated that scrutiny from social movements and empowered stakeholders in the country of origin are a clear mitigating factor against selective disclosure (Marquis, Toffel and Zhou 2016). Data triangulation is therefore essential, as IntRA10 noted: "Checks and balances mechanisms [are] quite important." IntRA4 talked about a "reality check" carried out by consulting stakeholders, such as NGOs or trade unions: "You are very much dependant on NGOs actually going to the field and researching about that." However, as IntE3 pointed out, this level of scrutiny is dependent on rating agencies making "efforts to go out of the box … it is time-consuming; it has a cost factor … but it is the only fair and reliable method if you want to go down to the subject." The assessment of controversies also helped counterbalance the sometimes overly positive information provided by companies.

Another challenge we identify is quantification. The rating agencies tended to favour quantitative data as they were perceived as more objective, efficient and reliable, in addition to being expected by the market: "We live in a society that likes quantifiable data better than qualitative." (IntRA3). But as IntE4 said: "Gathering quantitative data about things such as decent work is not straightforward." Companies generally report decent work-related information in a qualitative manner, providing limited quantitative data. As a result, agencies struggled to quantify companies' intentions and results: "It is hard to measure. We try to put scores … we try to be as data-driven as possible in our approach but, for human capital, it is actually quite hard." (IntRA8). At the same time, this approach was questioned by some of the interviewees, who pointed to the risk of losing important information that did not easily fit the evaluation grid and could not be quantified.

A third challenge is evaluating results beyond intentions. Rating agencies assess companies' intentions, commitments and policies, but measuring the concrete implementation and outcomes of these policies on working conditions and labour rights remains challenging. In the literature, the gap has been referred to as "decoupling", a misalignment between policies, implementation and outcomes. Scholars have argued that decoupling becomes even more manifest in increasingly complex business models relying on cross-border forms of production and outsourcing (see, for example, Bromley and Powell 2012; Graafland and Smid 2019; Kuruvilla et al. 2020). The nature of this challenge was made clear by IntA1: "On the one hand, you have these public reports and, on the other, you have the scandals. So, the scandals you pick up. But is this really a decent company to work for? … Well, you know their policies, but do they really function well?". Controversies partially help to overcome the problem of decoupling, as they identify scandals only when the damage has already been done. Moreover, they rely heavily on the limited capacity of affected stakeholders to express their grievances, and of civil society actors to report on them, probably allowing many incidents to go unreported.

4.3.2. The nature of decent work issues

ESG rating agencies face two main challenges related to the nature and characteristics of decent work issues. The first relates to contextualization. Companies do not exist in a vacuum. Rather, their practices are shaped by institutional context (DiMaggio and Powell 1983), including the normative framework in which they function, which defines the level of protection afforded to workers. The internationalization of businesses has exposed companies to a more diversified set of stakeholders and institutions. Studies have shown that geographical diversification and differences in stakeholders' power are key determinants of CSR performance (Abriata and Delautre 2020; Jackson and Rathert 2017). The literature has also pointed to the importance of strong local public regulations and empowered civil society stakeholders in bringing CSR commitments into effect in global supply chains, especially in the case of enabling rights, such as the freedom of association (Louche, Staelens and D'Haese 2020; Stroehle 2017). To evaluate decent work, agencies need to consider a multiplicity of standards defined at different levels of governance (at the global, transnational, sectorial, national and company levels) among different types of actors (public or private), subject to a variety of enforcement mechanisms (from purely voluntary to legally binding). This multiplicity raises the challenge of identifying the appropriate standards and methodologies to assess companies' practices on decent work in their varying operating environments. IntRA5 reflected, for example, on the varying degree of requirements concerning the management of ethnic diversity. In certain countries, companies have to report on ethnic diversity, while in others, companies are forbidden from collecting and reporting such information to guard against discriminatory practices. Beyond legislation, cultural and other societal aspects make comparative work more complex. For example, several interviewees observed that the use of labour courts or strikes in labour disputes was extremely variable from one country to another, making these indicators poor estimates of labour conditions from a comparative perspective. The ever-growing number of companies being rated is likely to make this issue even more critical.

The second challenge relates to objectivity. Most interviewees emphasized the need for ESG rating agencies to be objective, that is, "to eliminate subjectivity as much as possible" (IntE1). While there is no formal definition of objectivity in the ESG literature, the accounting literature regards it as a criterion of reliability; in other words, "the degree of closeness to being right" (Ijiri and Jaedicke 1966, 479). In the past, some agencies were criticized for their lack of objectivity due to their reliance on analysts' interpretations (Chatterji and Levine 2006; Berg, Kölbel and Rigobon 2022; Fowler and Hope 2007). Objectivity provides important proof of professionalism and credibility (Stubbs and Rogers 2013). As IntRA10 noted, "we are not a research firm, we don't provide subjective opinions". Accordingly, interviewees mentioned the need to provide "neutral" evaluations (IntE4), without a "moral angle or moral consideration" (IntRA2), based on "factual and objective data" (IntRA10) and quantitative data: "ideally, we would like to have 'right or wrong' – we try to have quantitative models behind our scores. At least, that suggests a level of objectivity." (IntRA8). Another important way of showing objectivity is through an audit trail – a system that traces data sources. The use

of publicly available information was presented as essential in documenting decisions: "Only publicly available data is used in our processes … because that is the only way we can show auditability and transparency" (IntRA10). However, as discussed earlier, public information tends to provide an overwhelmingly positive image of companies, which encourages agencies to consult different stakeholders in order to obtain more reliable information. As underlined by IntRA6: "It is really important to talk to someone from a trade union … he has experience, he will immediately know the ten best companies in his sector in terms of decent work circumstances". However, this practice did not seem to be systematized: "the intention was indeed to approach all stakeholders … but in reality, it was far less" (IntRA6).

Interviewees recognized that a certain level of subjectivity was inevitable when evaluating decent work. They regularly faced ethical dilemmas, such as where to set the boundary between good and bad practices, how to define scales for the severity of controversies or the diligence of companies in ensuring good working conditions: "there is no limit to [the] quality or quantity of social impact … how many jobs should a company provide? What is the right level of employment for a company? Eh, it is an almost impossible question." (IntA3). In contrast to the environmental dimension, ethical dilemmas seem to be more significant when assessing the effect of decent work on human relations. Moreover, owing to the concentration of agencies in Europe and North America, some interviewees considered issues related to decent work to be politicized, possibly favouring a "Eurocentric or Western-centric notion of what is good" (IntRA6).

4.3.3. The nature of ESG rating agencies

The nature of agencies relates to their organizational form and the characteristics of the ESG data market. We identify two challenges in this regard. The first concerns the principle of materiality, which, as explained earlier, focuses on what really matters in business terms and reduces the amount of unnecessary information. In the context of ESG assessment, the concept of materiality is theoretically expanded by considering a broader set of stakeholders (Global Reporting Initiative 2013). The agencies' most important customers were financial actors – asset managers, financial analysts and investors – working according to a shareholder value logic. Therefore, assessing decent work was generally done "from an investment perspective … what define[s] the risks and opportunities for companies" (IntRA2). Based on this logic, if an issue was not considered financially material, it would be regarded as irrelevant for investment. That does not necessarily mean that the issue was not relevant for decent work, but simply that there was no obvious link with the firm's financial value. This highlights a tension about what matters: "Some things, you can decide on the financial materiality argument, but other things, like human rights, we don't expect that you have to make the choice based on risks and returns" (IntRA2). The previous section showed how this tension can lead to very different technical arrangements between agencies concerning, for example, the inclusion of normative criteria in the weighting of ESG indicators or in the rating of controversies. The materiality principle tends to lock rating agencies into a short-term perspective, whereas the very definition of sustainability refers to organizations' ability to

balance the short and long terms. Some interviewees highlighted this contradiction and its importance with regard to decent work: “issues related to decent work might take more time to implement” (IntE4).

The second challenge relates to the market logic. Rating agencies are for-profit companies and face fierce competition. They need to offer a certain level of added value in order to remain competitive (“I think, as an investor, one has to be careful to select the company or the rating provider that aligns best with one’s own values in order to have an assessment that is meaningful for oneself.” IntRA3) and to be efficient in evaluating an ever-increasing number of companies. As table 1 shows, the ESG rating agencies in our study monitored between 4,700 and 11,000 companies. Yet, the lack of quantitative and clear data on decent work undermined efficiency and made evaluations lengthier. An interviewee indicated that: “rating agencies do not have the capacity [in terms of time] to look deeply into these things” (IntRA1). Besides, so far, the customers of rating agencies have shown limited interest in, or little mobilization for, decent work issues, in contrast to their reactions to other concerns like climate change: “it is very seldom that they [the clients] demand concrete data on labour rights” (IntRA4). Interviewees noted that greater demand from clients would be likely to encourage rating agencies to develop more sophisticated methods to assess decent work: “if there would be a demand for more concrete data on labour rights, the rating agencies would react to that, then maybe they would improve the criteria” (IntRA4).

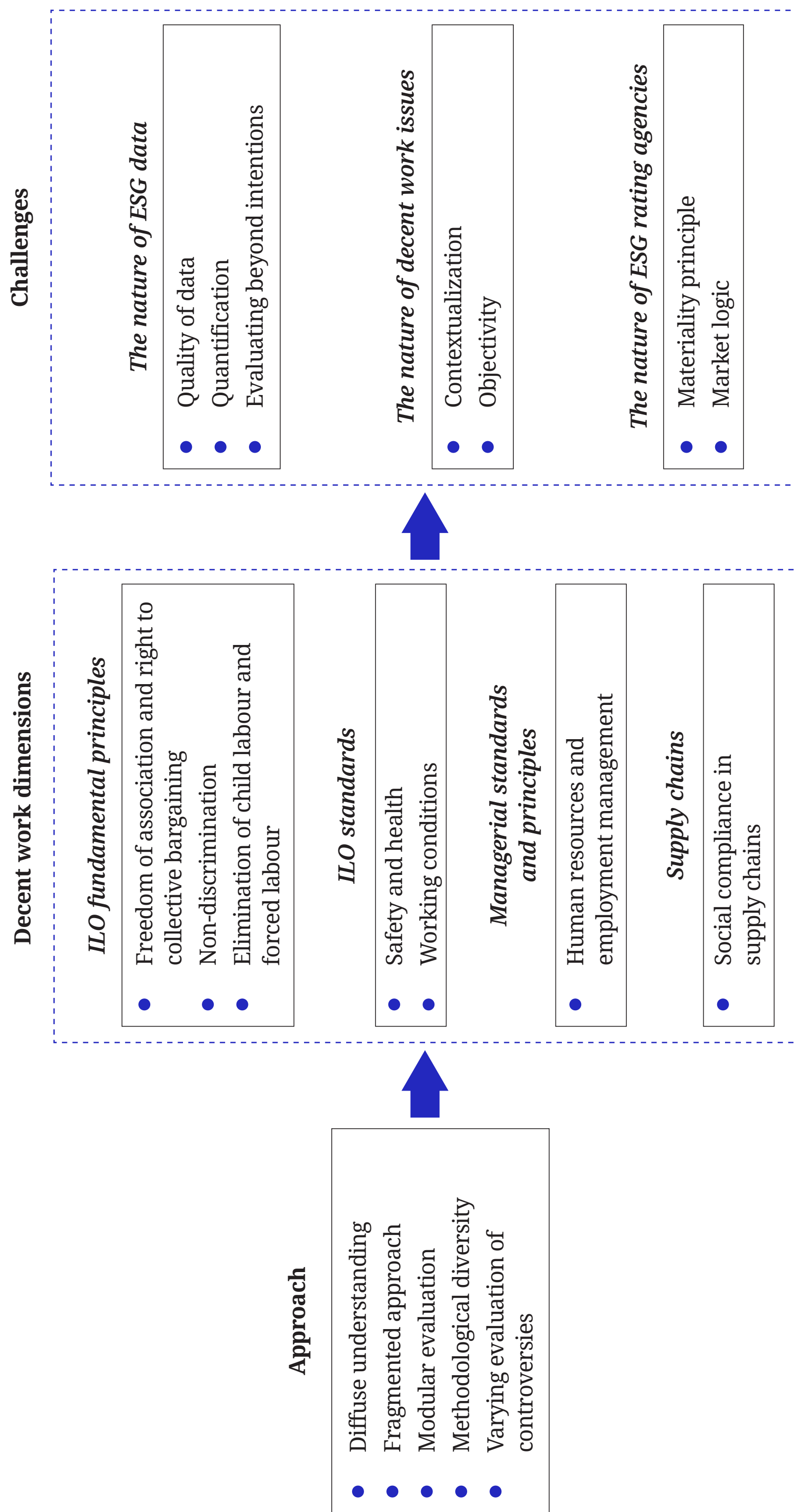
5. Discussion and conclusions

ESG rating agencies are key actors in assessing the CSR performance of companies. In this study, we have explored how ESG rating agencies assess companies’ performance on decent work and the challenges they face. Figure 2 provides an overview of our results.

These results show that decent work is a well-established and unquestionable dimension of ESG ratings. However, the notion of decent work itself remains diffuse and abstract for many ESG professionals. Although all the agencies in our study made explicit reference to ILO Conventions and allocated a substantial share of their activity to measuring indicators related to decent work, none of them used the term “decent work” itself. They usually referred to specific elements, such as freedom of association, child labour, gender equality, health and safety and working conditions, rather than the general concept. As a result, those indicators were spread throughout various sections of evaluations, making it difficult to get a sense of companies’ overall approach to decent work.

This study highlights the variety of approaches and methods developed by ESG rating agencies to assess decent work, confirming results from previous studies (Berg, Kölbel and Rigobon 2022). This variation creates a lack of consistency and alignment among agencies (Chatterji et al. 2016; Dorfleitner, Halbritter and Nguyen 2015) and a source of confusion for users, be it investors, researchers or companies themselves (Berg, Kölbel and Rigobon 2022; Scalet and Kelly 2010). We have also indicated that there are differences regarding

Figure 2. Challenges in the assessment of decent work



Source: Authors' compilation.

the scope of assessment, both in terms of breadth (the number of indicators covering decent work) and depth (the level of details with which each indicator is evaluated), and in methodologies for score aggregation, weighting, activation of indicators and the assessment of controversies. Moreover, our findings show that, although rating agencies try to provide the most objective evaluations possible, some degree of subjectivity is not only unavoidable but also desirable when assessing decent work.

Evaluating and measuring decent work remains a cumbersome task for ESG rating professionals. A central problem is the lack of availability, reliability and comparability of data (Chatterji et al. 2016; Christensen, Serafeim and Sikochi 2022; Widyawati 2021). As previous studies have shown, non-financial reporting differs significantly across regions and sectors (Fifka 2013; Ali, Frynas and Mahmood 2017). Interviewees in our study also highlighted the discrepancy between policies and outcomes. Companies engage in what scholars have identified as decoupling practices, through which they favour symbolic commitments over concrete measures (Bromley and Powell 2012; Tashman, Marano and Kostova 2019). Recent research has shown that the increasing complexity of business models is a source of field opacity for lead companies and can generate more decoupling (Kuruvilla, Li and Jackson 2021; Wijen 2014). Other studies have shown that formulating policies may actually generate a sense of entitlement in the company and a productivity narrative, which may trigger action (Haack, Schoeneborn and Wickert 2012; Zeffane, Polonsky and Medley 1994). Formulating policies is therefore a significant starting point in implementing decent work practices. Still, overcoming the challenge of decoupling is not easy, although it could be done through site visits and in-depth and continuous dialogue with stakeholders. However, such an approach would involve significant additional investment in terms of costs and time, which would affect agencies' business models. This also raises the question of the role of ESG rating agencies. As an interviewee stated: "We are not investigators. … We are here to provide a reasonable picture on the capacity of enterprises, on one side, to respect the principles and objectives defined by international Conventions and, on the other side, to protect the reputational capital" (IntRA1). Cooperating as a network with grassroot organizations (Goodman et al. 2014) and decent work experts could give analysts an external perspective with which to triangulate the information provided by companies, thereby promoting the design of indicators that go beyond policy commitments. Although agencies work in partnerships with organizations such as NGOs, collaborations remain limited. Using collaborative platforms to share information and involve multiple actors, including the rating agencies themselves, could be a powerful tool for assessing, and engaging with, companies on issues related to decent work.

Our study has also shown that analysts tend to favour quantifiable information, perceived as more objective and efficient, and information on issues that are (financially) material. We expect this trend to grow in the future owing to the recent move to market consolidation. However, quantitative data on decent work are scarce, which seriously limits what can be measured. Moreover, not all decent work issues are material in a financial sense or when considered from a short-term perspective. Such a strong material focus, and pre-formatted rating

grids, could lead to a narrow outlook on decent work and a loss of nuance in evaluation, highlighting the importance of contextualizing information. Quantified and material procedures thus show their limits when considering the complex regulatory framework in which companies operate at the transnational level.

In the light of those challenges, it is legitimate to wonder about the capacity of ESG rating agencies (and therefore RI) to address decent work, even if this article does not call into question the principle of rating itself. Agencies play an influential and important role in improving non-financial information disclosure. As an interviewee observed, they are "an accelerator for transparency" (IntRA7). Notably, some investor-led initiatives focusing on decent work, such as the Workforce Disclosure Initiative, the Corporate Human Rights Benchmark and working groups within the PRI, seek to improve the disclosure of information on decent work and dialogue with companies. Moreover, by acquiring ESG rating agencies, mainstream credit rating agencies like Moody's, MSCI, S&P Global and Morningstar have contributed to making ESG data more visible, thereby increasing the capacity of ESG ratings to influence companies. However, agencies remain limited in their capacity to capture the quality of decent work within companies. They are constrained by their own methodologies and the data available. They assess companies' intentions but can only partially evaluate policy outcomes. Controversies provide a necessary mechanism to counterbalance methodological limitations, but companies are unfortunately not equally targeted by the media and watch groups. Furthermore, if an event emerges as a controversy, it often means that serious damage has already been done.

Our study has highlighted the importance of active engagement by ESG data users, namely by establishing an open dialogue with rating agencies, acting carefully when considering ESG data and understanding the logic and approach of raters. However, there are several obstacles for such a proactive use of ESG information. First, agencies should be transparent about their underlying assumptions, methodologies and data sources. As our study has shown, this is not always the case. An interesting area for future research would be to study the impact of field consolidation on the quality and transparency of the ratings, especially in terms of their social dimension, which is subject to more variety than other dimensions. Second, Amel-Zadeh and Serafeim's (2018) survey showed that 82 per cent of investors use ESG information because it is financially material to investment performance. They are therefore primarily driven by financial rather than ethical motives, suggesting a limited interest in actual ESG issues. This questions the willingness of investors to be active users and their capacity to exert pressure on rating agencies to be more transparent and refine their indicators. For that to happen, other ESG rating agencies should seek to attract other users, such as trade unions or human rights organizations. A diversity of users would encourage agencies to move beyond the financial logic of ESG information and require them to consider other dimensions in their evaluations.

Lastly, it is also important to consider ESG rating agencies as part of a complex ecosystem where action from different quarters is likely to be mutually reinforcing for the promotion of decent work. The quality and reliability of ESG ratings depend largely on the ongoing movement supported by public and private actors to improve corporate reporting and transparency. Losing the

momentum would inevitably have negative effects on the quality of ratings. Moreover, the capacity of ESG ratings (and, more widely, of the RI movement) to address decent work depends on the involvement of a series of actors (namely, investors, companies and other stakeholders, including trade unions and other civil society groups) and their joint influence in reappraising materiality to better account for the complexities and the long-term perspective of decent work issues. In this light, future research could analyse the inclusion of decent work issues in investors' strategies, and study the role and influence of shareholder activism in this regard. It would also be relevant to explore how companies themselves can shape ESG ratings. Scholars have explored how companies respond to ESG ratings (Slager and Gond 2022), but not yet how they influence them. Some companies are doing pioneering work on decent work without necessarily being well-evaluated or recognized by rating agencies, owing to a mismatch between the evaluation criteria and their forward-looking approach.

More globally, and in follow-up to the ILO Centenary Declaration for the Future of Work, 2019, which promoted policies and incentives that allow a better alignment of business practices with decent work, it would be interesting to question the role of ILO constituents in structuring the RI and ESG rating markets. In addition to their actions to improve the transparency of business actors on extra-financial issues, governments also have a role to play in regulating financial actors – especially when those actors manage funds supplied by employees' savings for social protection purposes.[7] Similarly, the impact of action by social partners could also be a subject for future research. As mentioned above, certain agencies use trade unions as providers of extra-financial information to supplement other sources and, according to Penalva-Icher (2008), in addition to their traditional roles, unions in many countries can represent staff interests in the management of their savings plans. In theory, these areas of intervention could give unions leverage within a shareholder activism strategy. For their part, employers can be vocal critics of the absence of standardization in the area of ESG rating (MEDEF, Afep, Cliff and C3D 2019) at the national level. However, their interest in ensuring a level playing field for the assessment of companies' decent work policies and practices should also encourage them to consider the issue from a more global perspective.

References

Abriata, Bruno Dante, and Guillaume Delautre. 2020. "What Drives CSR? An Empirical Analysis on the Labour Dimensions of CSR", ILO Working Paper No. 7. Geneva: ILO.

Ali, Waris, Jedrzej George Frynas, and Zeeshan Mahmood. 2017. "Determinants of Corporate Social Responsibility (CSR) Disclosure in Developed and Developing Countries: A Literature Review". *Corporate Social Responsibility and Environmental Management* 24 (4): 273–294.

Amel-Zadeh, Amir, and George Serafeim. 2018. "Why and How Investors Use ESG Information: Evidence from a Global Survey". *Financial Analysts Journal* 74 (3): 87–103.

[7] For example, in the United States there is an ongoing debate concerning the recent statement from the US Department of Labor on its rules on ESG investments for employees' retirement plans (United States, Department of Labor 2021). See also Penalva-Icher (2008) on the case of France.

Avetisyan, Emma, and Kai Hockerts. 2017. "The Consolidation of the ESG Rating Industry as an Enactment of Institutional Retrogression". *Business Strategy and the Environment* 26 (3): 316–330.

Berg, Florian, Julian F. Kölbel, and Roberto Rigobon. 2022. "Aggregate Confusion: The Divergence of ESG Ratings". *Review of Finance* 26 (6): 1315–1344.

Bourguignon, Rémi, and Arnaud Mias, eds. 2017. "Les accords-cadres internationaux: étude comparative des ACI conclus par les entreprises françaises". Report for the ILO Office for France. Paris: ILO, IAE Paris and Université Paris Dauphine.

Bouten, Lies, Charles H. Cho, Giovanna Michelon, and Robin W. Roberts. 2017. "CSR Performance Proxies in Large-Sample Studies: 'Umbrella Advocates', Construct Clarity and the 'Validity Police'", Working Paper. https://doi.org/10.2139/ssrn.3107182.

Bromley, Patricia, and Walter W. Powell. 2012. "From Smoke and Mirrors to Walking the Talk: Decoupling in the Contemporary World". *Academy of Management Annals* 6 (1): 483–530.

Brown, Barb, and Mike Wallace. 2018. *The ESG Ecosystem: Understanding the Dynamics of the Sustainability Ratings & Rankings Landscape*. Cleveland, OH: BrownFlynn.

Chatterji, Aaron K., and David I. Levine. 2006. "Breaking Down the Wall of Codes: Evaluating Non-Financial Performance Measurement". *California Management Review* 48 (2): 29–51.

Chatterji, Aaron K., David I. Levine, and Michael W. Toffel. 2009. "How Well Do Social Ratings Actually Measure Corporate Social Responsibility?" *Journal of Economics & Management Strategy* 18 (1): 125–169.

Chatterji, Aaron K., Rodolphe Durand, David I. Levine, and Samuel Touboul. 2016. "Do Ratings of Firms Converge? Implications for Managers, Investors and Strategy Researchers". *Strategic Management Journal* 37 (8): 1597–1614.

Christensen, Dane M., George Serafeim, and Anywhere Sikochi. 2022. "Why Is Corporate Virtue in the Eye of the Beholder? The Case of ESG Ratings". *Accounting Review* 97 (1): 147–175.

Corbin, Juliet, and Anselm Strauss. 2008. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. 3rd ed. Thousand Oaks, CA: Sage.

Creswell, John W. 1998. *Qualitative Inquiry and Research Design: Choosing among Five Traditions*. Thousand Oaks, CA: Sage.

———. 2013. *Qualitative Inquiry and Research Design: Choosing among Five Approaches*. 3rd ed. Thousand Oaks, CA: Sage.

Delmas, Magali A., Dror Etzion, and Nicholas Nairn-Birch. 2013. "Triangulating Environmental Performance: What Do Corporate Social Responsibility Ratings Really Capture?" *Academy of Management Perspectives* 27 (3): 255–267.

DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields". *American Sociological Review* 48 (2): 147–160.

Dimmelmeier, Andreas. 2020. "Mergers and Acquisitions of ESG Firms: Towards a New Financial Infrastructure?" SocArXiv Working Paper, 13 October 2020. https://doi.org/10.31235/osf.io/jt2uk.

Dorfleitner, Gregor, Gerhard Halbritter, and Mai Nguyen. 2015. "Measuring the Level and Risk of Corporate Responsibility: An Empirical Comparison of Different ESG Rating Approaches". *Journal of Asset Management* 16 (7): 450–466.

Dumas, Christel, and Céline Louche. 2016. "Collective Beliefs on Responsible Investment". *Business & Society* 55 (3): 427–457.

Eccles, Robert G., and George Serafeim. 2013. "The Performance Frontier: Innovating for a Sustainable Strategy". *Harvard Business Review*, May 2013. https://hbr.org/2013/05/the-performance-frontier-innovating-for-a-sustainable-strategy.

Eccles, Robert G., and Judith C. Stroehle. 2018. "Exploring Social Origins in the Construction of ESG Measures", Working Paper. Oxford: University of Oxford, Saïd Business School. https://dx.doi.org/10.2139/ssrn.3212685.

Eccles, Robert G., Linda-Eling Lee, and Judith C. Stroehle. 2020. "The Social Origins of ESG: An Analysis of Innovest and KLD". *Organization & Environment* 33 (4): 575–596.

Fifka, Matthias S. 2013. "Corporate Responsibility Reporting and Its Determinants in Comparative Perspective: A Review of the Empirical Literature and a Meta-analysis". *Business Strategy and the Environment* 22 (1): 1–35.

Fortanier, Fabienne, Ans Kolk, and Jonatan Pinkse. 2011. "Harmonization in CSR Reporting: MNEs and Global CSR Standards". *Management International Review* 51 (5): Article No. 665. https://doi.org/10.1007/s11575-011-0089-9.

Fowler, Stephen J., and Chris Hope. 2007. "A Critical Review of Sustainable Business Indices and their Impact". *Journal of Business Ethics* 76 (3): 243–252.

Gioia, Dennis A., Kevin G. Corley, and Aimee L. Hamilton. 2013. "Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology". *Organizational Research Methods* 16 (1): 15–31.

Global Reporting Initiative. 2013. *G4 Sustainability Reporting Guidelines: Implementation Manual*. Amsterdam.

Global Sustainable Investment Alliance. 2021. *Global Sustainable Investment Review 2020*. http://www.gsi-alliance.org/wp-content/uploads/2021/08/GSIR-20201.pdf.

Golafshani, Nahid. 2003. "Understanding Reliability and Validity in Qualitative Research". *Qualitative Report* 8 (4): 597–606.

Goodman, Jennifer, Céline Louche, Katinka C. van Cranenburgh, and Daniel Arenas. 2014. "Social Shareholder Engagement: The Dynamics of Voice and Exit". *Journal of Business Ethics* 125 (2): 193–210.

Graafland, Johan, and Hugo Smid. 2019. "Decoupling among CSR Policies, Programs, and Impacts: An Empirical Study". *Business & Society* 58 (2): 231–267. https://journals.sagepub.com/doi/10.1177/0007650316647951.

Haack, Patrick, Dennis Schoeneborn, and Christopher Wickert. 2012. "Talking the Talk, Moral Entrapment, Creeping Commitment? Exploring Narrative Dynamics in Corporate Responsibility Standardization". *Organization Studies* 33 (5–6): 815–845.

Hoogervorst, Hans. 2017. "Making Sense of Materiality". *IFAC*, 25 September 2017. https://www.ifac.org/knowledge-gateway/supporting-international-standards/discussion/making-sense-materiality.

Ijiri, Yuji, and Robert K. Jaedicke. 1966. "Reliability and Objectivity of Accounting Measurements". *Accounting Review* 41 (3): 474–483.

ILO. 1999. *Decent Work*. Report of the Director-General, International Labour Conference, 87th Session, 1999. Geneva.

———. 2017. *World Social Protection Report 2017–19: Universal Social Protection to Achieve the Sustainable Development Goals*. Geneva.

Jackson, Gregory, and Nikolas Rathert. 2017. "Private Governance as Regulatory Substitute or Complement? A Comparative Institutional Approach to CSR Adoption by Multinational Corporations". In *Multinational Corporations and Organization Theory: Post Millennium Perspectives*, edited by Christoph Dörrenbächer and Mike Geppert, 445–478. Bingley: Emerald.

Kölbel, Julian F., Florian Heeb, Falko Paetzold, and Timo Busch. 2020. "Can Sustainable Investing Save the World? Reviewing the Mechanisms of Investor Impact". *Organization & Environment* 33 (4): 554–574.

Kuruvilla, Sarosh, Mingwei Liu, Chunyun Li, and Wansi Chen. 2020. "Field Opacity and Practice–Outcome Decoupling: Private Regulation of Labor Standards in Global Supply Chains". *Industrial and Labour Relations Review* 73 (4): 841–872.

Kuruvilla, Sarosh, Ning Li, and J. Lowell Jackson. 2021. "Private Regulation of Labour Standards in Global Supply Chains: Current Status and Future Directions". In *Decent Work in a Globalized Economy: Lessons from Public and Private Initiatives*, edited by Guillaume Delautre, Elizabeth Echeverría Manrique and Colin Fenwick, 185–210. Geneva: ILO.

Louche, Céline, Lotte Staelens, and Marijke D'Haese. 2020. "When Workplace Unionism in Global Value Chains Does Not Function Well: Exploring the Impediments". *Journal of Business Ethics* 162 (2): 379–398.

Lucchini, Andrea, and Anna Maria Moisello. 2017. "CSR Disclosure, Visibility and Media Pressure: International Evidence from the Apparel and Textile Industry". *European Journal of Economics, Finance and Administrative Sciences* 93 (February): 5–29.

Marquis, Christopher, Michael W. Toffel, and Yanhua Zhou. 2016. "Scrutiny, Norms, and Selective Disclosure: A Global Study of Greenwashing". *Organization Science* 27 (2): 483–504.

MEDEF (Mouvement des Entreprises de France), Afep (Association française des entreprises privées), Cliff Investor Relations (Association française des professionnels de la communication financière) and C3D (Collège des Directeurs du Développement Durable). 2019. "French Initiative on the Relations between Companies and Non-financial Rating Agencies". February 2019. http://www.afep.com/wp-content/uploads/2019/02/Afep-Medef-Cliff-C3D-initiative-on-non-financial-rating-February-2019.pdf.

Nauman, Billy. 2019. "Credit Rating Agencies Join Battle for ESG Supremacy". *Financial Times*, 17 September 2019. https://www.ft.com/content/59f60306-d671-11e9-8367-807ebd53ab77.

Novethic. 2018. "SDGs: A New Focus for Non-Financial Rating Agencies". September 2018. Paris.

Penalva-Icher, Élise. 2008. "Comment devenir légitimes? Le double rôle des syndicats dans le marché de l'investissement socialement responsable". *La Revue de l'Ires* 57 (2): 111–140.

———. 2016. "La professionnalisation dans l'Investissement socialement responsable. Le cas des analystes extra-financiers". *Revue Française de Socio-Économie* 2016/1, No. 16: 141–159.

PRI (Principles for Responsible Investment). 2018. "Primer on responsible investment in infrastructure: Introduction". https://www.unpri.org/infrastructure/primer-on-responsible-investment-in-infrastructure-/2700.article#:~:text=The%20Principles%20for%20Responsible%20Investment,%2C%20long%2Dterm%20returns%E2%80%9D.

Scalet, Steven, and Thomas F. Kelly. 2010. "CSR Rating Agencies: What Is Their Global Impact?" *Journal of Business Ethics* 94 (1): 69–88.

Sekerler Richiardi, Pelin, and Ma. Diyina Gem Arbo. 2019. "Multinational Enterprises and Social Protection: A Case Study of the L'Oréal Share & Care Program", Research Department Working Paper No. 42. Geneva: ILO.

Semenova, Natalia, and Lars G. Hassel. 2015. "On the Validity of Environmental Performance Metrics". *Journal of Business Ethics* 132 (2): 249–258.

Sjöström, Emma. 2008. "Shareholder Activism for Corporate Social Responsibility: What Do We Know?" *Sustainable Development* 16 (3): 141–154.

Slager, Rieneke, and Jean-Pascal Gond. 2022. "The Politics of Reactivity: Ambivalence in Corporate Responses to Corporate Social Responsibility Ratings". *Organization Studies* 43 (1): 59–80.

Stroehle, Judith Christina. 2017. "The Enforcement of Diverse Labour Standards through Private Governance: An Assessment". *Transfer: European Review of Labour and Research* 23 (4): 475–493.

———. 2018. "The Social Origins of ESG? An Analysis of Innovest and KLD". Paper presented at the Yale Initiative on Sustainable Finance Symposium 2018: State of ESG Disclosure Standards, New Haven, CT, 2 November 2018.

Stubbs, Wendy, and Paul Rogers. 2013. "Lifting the Veil on Environment-Social-Governance Rating Methods". *Social Responsibility Journal* 9 (4): 622–640.

Tashman, Peter, Valentina Marano, and Tatiana Kostova. 2019. "Walking the Walk or Talking the Talk? Corporate Social Responsibility Decoupling in Emerging Market Multinationals". *Journal of International Business Studies* 50 (2): 153–171.

United States, Department of Labor. 2021. "US Department of Labor Releases Statement on Enforcement of Its Final Rules on ESG Investments, Proxy Voting by Employee Benefit Plans", 10 March 2021. https://www.dol.gov/newsroom/releases/ebsa/ebsa20210310.

Widyawati, Luluk. 2021. "Measurement Concerns and Agreement of Environmental Social Governance Ratings". *Accounting & Finance* 61 (S1): 1589–1623.

Wijen, Frank. 2014. "Means versus Ends in Opaque Institutional Fields: Trading Off Compliance and Achievement in Sustainability Standard Adoption". *Academy of Management Review* 39 (3): 302–323.

Zarlowski, Philippe. 2007. "Marketing Corporate Social Responsibility in a National Context: The Case of Social Rating Agencies in France". In *Managing Corporate Social Responsibility in Action: Talking, Doing and Measuring*, edited by Frank den Hond, Frank G.A. de Bakker and Peter Neergaard, 167–186. Aldershot: Ashgate.

Zeffane, Rachid M., Michael J. Polonsky, and Patrick Medley. 1994. "Corporate Environmental Commitment: Developing the Operational Concept". *Business Strategy and the Environment* 3 (4): 17–28.