



HAL
open science

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

Céline Louche, Guillaume Delautre, Gabriela Balvedi Pimentel

► To cite this version:

Céline Louche, Guillaume Delautre, Gabriela Balvedi Pimentel. Assessing companies' decent work practices: An analysis of ESG rating methodologies. *International Labour Review*, 2022, 162 (1), pp.69-97. 10.1111/ilr.12370 . hal-03690210

HAL Id: hal-03690210

<https://audencia.hal.science/hal-03690210>

Submitted on 9 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Assessing companies' practices on decent work: an analysis of

ESG rating methodologies

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

*Audencia Business School, email: clouche@audencia.com; **ILO, email: delautre@ilo.org (corresponding author); ***consultant, email: gabrielabpimentel@gmail.com. This research benefited from the financial support of the French government as part of the cooperation agreement between the French government and the ILO. The authors are also grateful to Andrew Crane, Jake Reynolds and Fernando Messineo Libano for their helpful and constructive comments on earlier drafts of the paper.

Abstract

Environmental, social and governance (ESG) rating agencies, as non-financial data providers, have become a central actor 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 paper investigates how these actors measure and assess companies on decent work-related items and identifies the challenges they face in this endeavour. The paper aims to better understand the capacity of responsible investment and ESG ratings in promoting and improving decent work within companies.

Keywords

ESG rating, responsible investment, decent work, measurement, corporate social responsibility

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/ilr.12370](https://doi.org/10.1111/ilr.12370).

Decent work is increasingly considered an essential part of business responsibility. Promoted by the International Labour Organization (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 set up by the United Nations General Assembly in 2015 (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 legal such as the 2015 UK Modern Slavery Act, the 2017 French 'Duty of Care' Act or the EU directive 2014/95/EU, and voluntary, such as the Global Reporting Initiative or the Global Compact. All these initiatives require companies to publicly disclose non-financial information, including on decent work.

In this context, evaluating companies' commitments to and practices of decent work is central. The question became even more salient with 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 2020). ESG related data, including decent work, are essentially provided by ESG rating agencies. In the past twenty years, agencies have attracted scholars' attention, who have studied their role as metric providers (Chatterji, Levine, and Toffel 2009), their methodologies (Eccles, Lee, and Strohle 2020, Eccles and Strohle 2018, Penalva-Icher 2016, Strohle 2019) as well as measurement challenges (Berg, Kölbel, and Rigobon 2020, Chatterji et al. 2016). Although these studies have provided rich insights, little is known about the specificities of each ESG dimension, especially decent work.

Accordingly, this paper 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? The objective is to get a better appreciation of RI and agencies' capacities in promoting decent work.

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

1. Responsible investment and ESG rating agencies

1.1. Responsible investment

Once a niche , RI has become part of mainstream financial markets (Dumas and Louche 2016). The 2020 Global Sustainable Investment Review estimated the RI market at over \$35 trillion (GSIA 2021). The United Nations Principles for Responsible Investment (UNPRI) counted more than 3,826 signatories in 2021, including the largest institutional investors and asset managers worldwide, representing over \$121 trillion under management¹.

¹ PRI website, accessed on May 07, 2022

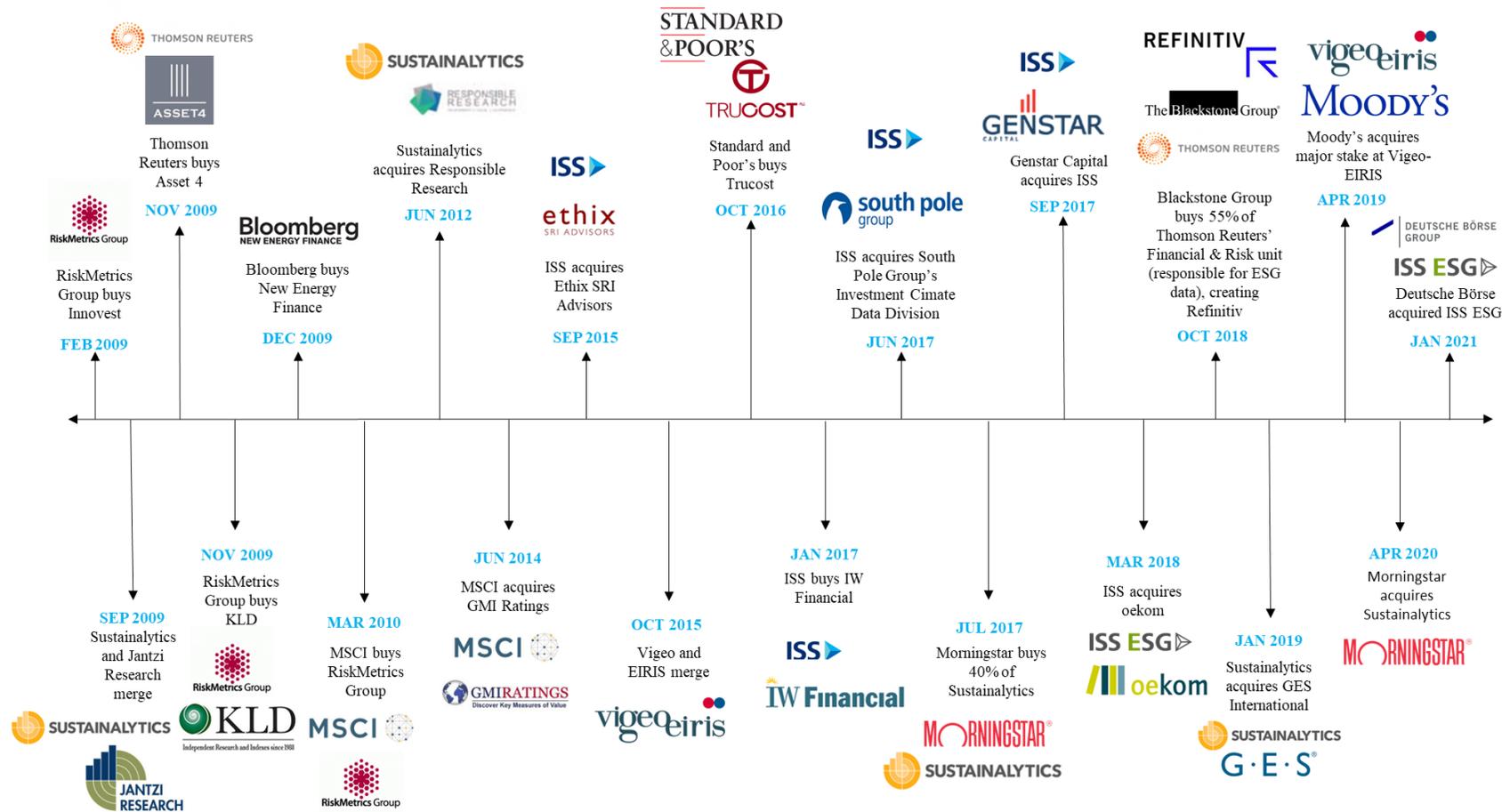
Although RI takes many forms and approaches, five main strategies have been identified in the literature: avoidance (the use of exclusionary criteria to avoid businesses or practices regarded as unacceptable or harmful to society such as alcohol, tobacco industries, involvement in human rights violations);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

1.2. ESG rating agencies

ESG rating agencies assess and compare companies on their CSR performance by collecting, aggregating and interpreting large quantities of information. They became prominent actors in the RI field and play a central role in constructing and operationalising ESG data (Zarlowski 2012). 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 past 30 years, the market has witnessed a strong consolidation trend 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). As a result, the market is experiencing increasing domination of American business actors as well as a convergence between mainstream financial actors and ESG actors (Novethic 2018, Nauman 2019).

Figure 1: Consolidation of the ESG rating agencies' market ²



² Adapted from Brown and Wallace (2018), p.6

1.3. ESG rating agencies as metrics providers

Agencies have designed in-house methodologies to assess companies' ESG performance. Several studies have investigated these methodologies and elaborated metrics. Empirical works have focused mainly on the aggregated ESG score (Chatterji et al. 2016, Chatterji, Levine, and Toffel 2009, Widyawati 2021) and the environmental dimension (Chatterji, Levine, and Toffel 2009, Semenova and Hassel 2015). Little has been done on the social dimension of ESG, or more precisely, on labour-related sub-dimensions.

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

The social origin of ESG rating agencies refers to the history of the rating organisation. It infuses the agency's understanding and conceptualisation of ESG and, consequently, how it is measured (Eccles, Lee, and Stroehle 2020). The approach taken to evaluate companies, including the type and construction of indicators, is directly influenced by how an agency defines its own mission. Eccles and Stroehle (2018) identified two types of agencies. The first is defined as "value-driven" and focuses on ESG information which is financially relevant. In this case, ESG is defined according to its capacity to create corporate value and shareholder return. The second type is defined as "values-based" for which sustainability and societal impacts are dominant. 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 helps to explain the idiosyncratic characteristics linked to the methodological

decisions made by data providers (Stroehle 2019). Nevertheless, the consolidation of the industry suggest an increasing convergence towards a more value-driven approach (Bouten et al. 2017, Eccles, Lee, and Stroehle 2020).

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" (Eccles and Stroehle 2018, 9). The same concept is used in RI to assess which factors of ESG matter the most for investors (Eccles, Lee, and Stroehle 2020). Materiality is a highly debated and controversial concept as it implies a prioritisation of information considered relevant for investment decisions, working as a filter through which management values information (IFAC 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 the problem of divergence among the ESG ratings (Chatterji et al. 2016, Christensen, Serafeim, and Sikochi 2020, Delmas, Etzion, and Nairn-Birch 2013, Berg, Kölbel, and Rigobon 2020) and a low level of correlation within ESG ratings (Chatterji, Levine, and Toffel 2009, Kölbel et al. 2019). Chatterji et al. (2016) showed that commensurability (e.g. "how raters measure the same constructs" p. 1600) among rating agencies is low, even when adjusted for explicit differences in their theorisation (e.g. "the beliefs rates have about what being socially responsible means" p. 1599) of ESG. They concluded that most assessments provided by agencies potentially encompass high measurement errors. In line with previous findings, Berg, Kölbel, and Rigobon (2020) distinguished two primary sources of divergence among ESG ratings: the aggregation divergence, referring to the scope of issues considered and the weight associated with them;

and, more importantly, the measurement divergence, consisting of differences regarding the indicators used. At a more disaggregated level, studies show that the correlation of the score related to the social dimension (including labour and rights at work issues) is generally lower across rating agencies than the environmental dimension and can be negligible and even negative according to the pair of agencies considered (Berg, Kölbel, and Rigobon 2020, Dorfleitner, Halbritter, and Nguyen 2015). Such differences and inconsistencies are likely to lead to significantly diverse investment recommendations, create confusion, and even misinforming investors (Delmas, Etzion, and Nairn-Birch 2013). That is why several business actors and academics have called for a move towards more regulation and standardisation in ESG measurement (Berg, Kölbel, and Rigobon 2020, Medef-Afep-Cliff-C3D 2019).

2. Research method

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

2.1. Data collection

Our study is based on the analysis of six major ESG rating agencies (see Table 2 for more details on the agencies). Primary and secondary data were collected from various sources. We conducted 18 semi-structured interviews, ten with agencies' representatives and eight with experts working at or in relation with the ILO, universities, or ESG specialised organisations (see Table 1). Interviews included questions on 1) their organisation and their background, 2) their understanding of decent work, 3) its assessment and the challenges faced in evaluating labour-related items, 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 organisations such as the Global Compact, stock markets, Novethic, PRI as well as media sources.

Table 1: Interview overview

Category	Organisation	Function	Duration (minutes)
Decent work experts	ILO	Social Finance/Enterprises (2 persons)	71
		Enterprises/CSR team	60
Decent work and ESG experts	ILO, ICFTU, Ethibel	Member of the Ethibel Register Committee.	69
ESG Experts	University	Professor in Ethics and Economics	84
		Research Fellow	64
		Professor in Applied Economics	70
	Think Tank/University	Executive Director	68
	United Nations' Principles for Responsible Investment	Manager, Decent Work Programme	90
Rating agencies	ISS-ESG	Senior Associate, ESG Ratings	84
		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	

Table 2: ESG rating agencies*

ESG rating organization	Number of companies (worldwide)	Number of ESG Analysts (approx.)	Number of ESG indicators	Headquarter	Indexes
Sustainalytics (Morningstar)	11000	250 (out of 650 employees)	220 and 450 indicators	USA	Global Sustainability Signatories Index ; Jantzi Social Index Also subsidizes the production of the two FTSE Rutsell ESG Indexes: FTSE Developed ESG Index and FTSE Emerging ESG Index
VigeoEiris (Moody's)	8000	140 (out of 245 employees)	330 indicators	USA	8 Euronext indexes and Ethibel Sustainability Index
MSCI	8500	185 research analysts	56 indicators, 230 datapoints	USA	ESG Rating covers companies included in the following 22 MSCI indexes
Refinitiv	9000	150 research analysts	450 data points and 70 analytics	USA	-
RobecoSAM (S&P Global)	4700	No information	An average of 100 data points and 23 criteria per company	USA	Used to inform the Dow Jones Sustainability Index
ISS ESG	8000 (aim to achieve 10000 by the end of 2020)	200 research analysts (out of 2000 employees)	An average of 100 criteria per company out of a pool of 800+ indicators	Germany	Supports the following indexes: 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

*Based on publicly available information from the agencies' websites and documents that were accessed between May and September

2020

2.2. Data analysis

Prior to the interviews, we organised our secondary data to gain an in-depth understanding of the selected agencies and decent work-related frameworks. We built tables to list the criteria, methodologies, and sources of information, which were later completed by information gathered through the interviews. This first step helped us to understand the way decent work was assessed and compare the agencies.

The second step consisted in analysing the interviews. Using NVivo, we applied a comprehensive coding procedure (Creswell 2013, Corbin and Strauss 2007). In support of our inductive approach and to ensure an open-minded analysis, the first initial coding of the data was based on open codes. We then moved back and forth between data and emerging concepts to finally 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 on the challenges linked to ESG data. The axial coding was done 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).

3. Results

Measuring decent work is an essential part of ESG ratings. All agencies, without exception, assess companies' policies and practices in this domain. The 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 the supply chains, are considered by most agencies. We estimate that decent work-related items represented approximately 20 per cent to 35 per cent

of the overall ESG score. Although just an approximation, it signals the importance of decent work in ESG ratings. However, and as we will see later, our analysis reveals difficulties to assess and comprehend companies' global approach to decent work.

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

3.1. Variety of approaches to assess decent work

Diffused understanding. The analysed agencies integrate dimensions of decent work but do 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). They preferred to use labels which, according to them, are better understood by investors such as "human rights" or more business-oriented such as "human capital management", "human resources" or "employee engagement". An interviewee noted that "*decent work sounds a bit activisty*" (IntE4). However, they all referred to ILO standards and the UN Global Compact. An expert (IntA1) expressed the importance of "opposability" of the criteria to international norms and instruments, such as the Universal Declaration of Human Rights or ILO core conventions .

Fragmented approach. It is difficult to have a fine-grained understanding of the way ESG rating agencies measure decent work. Due to business confidentiality, limited information is available. But above all, the measurement of decent work is fragmented and not always visible. We found decent work items in categories such as human rights, human resources, business relationships or supply chain and those items were not always explicitly mentioned or

sometimes hidden in subcategories. Also, multiple labels were used among different agencies to refer to the same dimensions of decent work.

Modular evaluation. Within a same agency, the number of items used to assess decent work varied substantially between the companies assessed but also among the different aspects evaluated within the same company. The number of criteria used changed depending on the sector and the region. Based on the perceived risks and relevance, indicators can be activated or deactivated. 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. We use criteria that are industry-specific*". But also within the same company, an indicator may be activated when assessing one part of its business and deactivated when assessing another. For example, an interviewee (IntRA3) informed 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.

Different weighting systems. Agencies apply a weighting system to the indicators, attributing varying weights to the indicators depending on in-house rules which can be related to sectors, locations, or other characteristics. For example, IntRA7 said that "*social issues are more highly weighted in industries where it is more service oriented*". Another interviewee mentioned the example of Human Resources 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, the details of weighting systems were not always made explicit by rating agencies.

The underlying approach to materiality, which is directly influenced by the social origin of the rating agencies, also influences agencies' weighting system. For example, one agency stipulated that the weight of each indicator at the industry level was based on its past impact on financial performance over time (i.e. measured by the level of correlation between the two measurements). Two other agencies included in their weighting system the risk exposure related to, for instance, companies' business models, products, geographical position, size, or reliance on public contracts or outsourced production. Those agencies were more value-driven and thereby put more weight on indicators that have directly impact 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 such as the nature of risks (e.g. if connected to universally agreed human rights) or the specific vulnerabilities of company's stakeholders (because it operates in particularly risky environments for workers' rights and working conditions) were considered alongside more financially related criteria such as reputation, human capital, operational efficiency or legal security.

Varying evaluation of controversies. An ESG controversy is an information of public nature which places the company under media spotlight and involves its responsibility on one or several rating items. Controversies usually refer to past or ongoing scandals such as cases of toxic waste spill, human rights violations, or corruption. ESG rating agencies identify controversies by continuously researching multiple media sources and NGOs' publications.

While all agencies considered controversies, their scoring methods varied significantly as they build on several layers of interpretation. Controversies were assessed based on their level of severity, which depends on the definition and understanding of the incident itself ("*Is this something local or global? Does it involve top management?*" (IntRA7), "*is there death of someone*" (IntRA9)) and the scale of the impact on the company (*was there a condemnation, a legal proceeding, or is it just an allegation?* (IntRA9)). They also considered the frequency of

controversies at the company level (i.e. is it a one-time event or a recurring pattern?). Finally, they considered dimensions such as the responsibility of the company (i.e. failure in monitoring, breach of company policies or legislations, court decisions, etc.), financially material criteria (fines, reputation, production stoppages) and the quality of company's responses to this controversy. The normative criterion emerged relatively marginal in this regard and were only considered by one value-based oriented agency. In this case, the criticality (e.g. is it related to a universally recognised right?), the amplitude (e.g. what is the number of stakeholders affected?) and the irreversibility of the controversies were considered.

3.2. Analysis of decent work items

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: freedom of association and right to collective bargaining, non-discrimination, child labour and forced labour. The following three refer to other ILO standards: health and safety, 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 the different categories including examples of indicators used by rating agencies to assess decent work.

Table 3: Examples of decent work items used by ESG rating agencies

Groups	Categories	Coverage	Scope	Evaluation mode	Examples of items
ILO Fundamental Rights	Freedom of association and collective bargaining	All agencies	Respect freedom of association and collective bargaining	Dominantly qualitative based on policy commitments, with some quantitative indicators	Guarantee 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
	Non-discrimination	All agencies	Mainly focused on diversity; limited consideration of other forms of diversity	Dominantly qualitative based on policy commitments, with some quantitative indicators	Trade union representation
					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
Other ILO Standards	Child labour and forced labour	All agencies (not always explicit)	Eliminate 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
					Health and Safety policy
	Occupational health and safety	All agencies	Commitments, implementation strategies, and certifications	Quantitative and qualitative	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
Working conditions	All agencies but not explicitly mentioned	Commitment to comply with applicable	Dominantly qualitative based on policy commitments, with some quantitative indicators	Judicial proceedings	
				Policy to ensure work-life balance	
				Working conditions policy	
				Quality of remuneration systems	
				Respect and management of working hours	
Employment security and types of employment					

	(except two agencies)	wage and hours legislation			
Human resources and employment management	All agencies	Commitment and implementation of career management, but also, for a few, restructurings 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
					Scope and quality of social supplier standards
Supply Chain	All agencies	Scope, means and coverage of suppliers' policies	Dominantly qualitative based on policy commitments, with some quantitative indicators		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 of sector specific frameworks
					Procedures to ensure compliance with labour rights

Freedom of association and right to collective bargaining is mentioned by all agencies, but their interpretation and measurement varied substantially. As an interviewee noted, this item's objective is to "*look at the extent to which the enterprise respects trade union freedom, collective bargaining rights and promotes collective bargaining right... but then what are the separate criteria?*" (IntA1). For most agencies, the indicators used and their scope, whether it applies to all rated companies or only to some, remained unclear. A minority of agencies considered this item to be part of the dimension dedicated to human rights, while most placed it within "labour management" or "human capital". Those items were generally measured on the basis of policy commitments, implemented actions, and monitoring and grievance mechanisms. Some agencies also considered companies' participation in the UN Global Compact or formalised bilateral commitments through International Framework Agreements (IFAs). Box 1 provides an illustration of how policy commitments are being evaluated by an agency. Certain agencies considered quantitative indicators such as the number of employees covered by independent trade unions or the coverage of collective bargaining agreements. Interestingly, two agencies asserted to have a specific focus on companies operating in countries with severe restrictions related to workers' collective rights. In these cases, companies were questioned on possible alternative workers participation measures which 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 labour relations was considered as a factor of risks for the business, leading to higher costs and lower profitability in the short term.

Box 1. Evaluation of companies' commitments to freedom of association

For this illustration, we build directly from the data of one of the analysed rating agencies. The agency measures a company's policy commitments with 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 enterprise.

Company A is a European company in the automotive industry. The agency has given the company the maximum score on policy commitment because: 1) the company has signed IFAs and other agreements with a major global union federation, specific agreements with unions in some countries of operation and the Global Compact; 2) the company's commitment undertaken in the IFA is detailed, addresses most of its responsibilities and explicitly supports implementation of ILO Conventions Nos. 87, 98, and 135; 3) the company's commitment applies at all levels of the firm, it is supported by senior management, and it collaborates with union representatives.

For a company to receive a low score in this domain, it should not disclose any public commitment, or not take part in any initiative.

Non-discrimination and more generally the promotion of diversity in enterprises is widely considered by rating agencies. However, agencies differ significantly regarding the degree of details to which they assess non-discrimination and the grounds of discrimination considered. In accordance with ILO Convention n° 111, half of the agencies considered multiple grounds of discrimination such as gender, ethnic origin, or age. The evaluation was mostly based on the commitments formalised by the company, such as discrimination policies, and the existence of concrete actions regarding training, recruitment or other management systems to fight discrimination. However, different grounds of discrimination are not equally considered. Indeed, gender equality and to a lesser extent disability are often given more importance than ethnicity or age. For some agencies, gender was even the only form of discrimination evaluated. The assessment of gender discrimination was more detailed and included quantitative indicators on the number of women employees, access to management, executive positions or pay gap.

Child and forced labour are less explicitly covered by agencies than the other ILO fundamental rights and principles. Only two agencies mentioned these items in their analytical grids, mostly through the evaluation of public commitments and instruments (e.g. code of conducts). In addition, these items seemed to only be activated for a companies with operations in countries or industries considered at-risk. Yet, the topic was usually considered in the evaluation of supply chains (see below) and was also monitored through controversies screening.

Health and safety issues are considered by all agencies and apply widely among different sectors. However, the type and number of indicators used to assess them differed per sector. Three of the agencies prioritised companies' policy commitments, implementation strategies, and certifications, such as OHSAS 18001 and ISO 45001. Another assessed health and safety by considering the number of judiciary procedures relating to this topic. Some of the agencies also reported to consider indicators on occupational injuries, illness, and fatalities. Interestingly, only one agency considered information linked to the prevention of stress-related diseases and mental health.

Other **working conditions** issues are considered although to a lesser extent. 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 (e.g. limiting working long hours and compensation rules for atypical or overtime). Another agency evaluated companies' compliance to applicable wage and hours legislations, and their adoption of industry-specific codes of conduct. Another was experimenting with fair wage, although it had not yet integrated this item in the 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 are universally applied.

In the myriad of disaggregated social items assessed by rating agencies, we identified a group of indicators related to **human resources and employment management**. Those items relate more specifically to career management, such as training systems, the existence of (individual) performance appraisal and talent identification mechanisms. Only half of the agencies gave a specific emphasis on employment security, mostly through indicators related to job turnover or the use of temporary employment. It should also be noted that two agencies evaluate the management of restructuring processes by analysing elements such as anticipation of reorganisation, involvement of employees' representatives or assistance provided to affected employees.

Social compliance in the management of the supply chains is widely recognised as an important topic, but with significant variations. Four agencies provided reasonable details on the way they evaluate decent work in the supply chain, including the scope, means and coverage of suppliers' policies. Agencies reported to consider a vast array of issues such as the four fundamental principles and rights at work, health and safety, (living) wages, working time, violence and harassment. The range of issues evaluated here was often vaster than the number of topics considered for the rated company's own employees. The agencies also aimed at evaluating the quality of existing procedures, such as social audits system, procedures in case of non-compliance, and internal training for purchase managers. Specific industry-related topics, for example conflict minerals policies and programmes, were the object of a special inquiry for some agencies. However, none of the agencies mentioned the level of the supply chain (the tiers) being considered

Finally, two decent work-related issues have attracted very little attention. The first is the **economic and social impacts of companies' investments and activities on local**

communities and businesses. The responsibilities of multinational companies in this regard have been recognised in the ILO Tripartite Declaration of Principles concerning Multinational Enterprises and Social Policy (see articles 17 to 20). However, this criterion was identified in the methodology of only one agency. The second is **social protection**, which is one of the four pillars of decent work. The issue was rarely mentioned in agencies' methodologies. While access to social protection is a prerogative of governments and public authorities in most countries, the scope, coverage and quality of social protection systems are highly uneven across the globe (ILO 2017). Therefore, a few pioneer MNEs have decided to directly provide social protection benefits and services such as death, accidents and disability insurance, paid maternity and paternity leave, healthcare or meal subsidies to their employees (see (Bourguignon and Mias 2017, Sekerler Richiardi and Arbo 2019) for concrete examples).

3.3. Challenges in measuring decent work

This section focuses on the structural, technical, and ethical challenges in measuring companies' performances on decent work. We organised them around three categories: the nature of ESG data, the nature of decent work issues and the nature of ESG rating agencies.

The nature of ESG data refers to the type of data required, and the methodologies used to assess decent work. We identified three closely inter-related challenges in this regard. The most important is the quality of the data, i.e., their reliability, comparability, and availability. ESG rating agencies rely prominently on publicly available information provided by rated companies. Research has shown that reporting practices among companies differ greatly both in terms of quantity and quality of the information depending on factors such as size, profitability, or sectors, but also the influence of institutional country-related factors including access to freedom of expression and media (Ali, Frynas, and Mahmood 2017, Marquis, Toffel, and Zhou 2016, Fifka 2013, Lucchini and Moisello 2017, Fortanier, Kolk, and Pinkse 2011). Although this challenge concerns all CSR dimensions, our interviewees noted that it was more

salient for 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 strings on the quality and the effective respect of labour rights"* (IntRA6). Although practices differed among agencies, in the case of absence of information, companies were usually downgraded: *"the lowest grade means that there is no information"* (IntRA3). When data was available, the question of reliability and exhaustiveness was raised. Many analysts mentioned the difficulty to interpret the provided information *"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 highlighted that scrutiny from social movements and empowered stakeholders in country of origin are a clear mitigating factor against selective disclosure (Marquis, Toffel, and Zhou 2016). Data triangulation is therefore essential as IntRA10 continued: *"checks and balances mechanisms is quite important"*. IntRA4 talked about *"reality check"* done through 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 said, it requires the rating agencies to put *"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 enabled counterbalancing the sometimes overly positive information provided by companies.

Another challenge is quantification. Rating agencies tend to favour quantitative data as it is perceived as more objective, efficient, and reliable and 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, with very limited provision of 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, such an approach was questioned by some of the interviewees, who pointed to the risk of losing important information that does not fit easily in the evaluation grid and cannot 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 as decoupling, a misalignment between policies, implementation and outcomes. Scholars have argued that decoupling becomes even more manifest with increasingly complex business models relying on cross-border forms of production and outsourcing (e.g. Bromley and Powell 2012, Graafland and Smid 2019, Kuruvilla et al. 2020). This difficulty is made clear in this quote: *"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?"* (IntA1). Controversies help to partially overcome the problem of decoupling, as it picks up scandals only when damages have already occurred. Moreover, it relies heavily on the capacity of affected stakeholders to express their grievance and of civil society actors to report on them, likely missing many incidents that go unreported.

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 contextualisation. Companies' practices do not happen in a vacuum but are shaped by the institutional context in which they operate (DiMaggio and Powell 1983), including the normative framework in which companies' function. It defines the level of protection afforded to workers. The internationalisation of businesses exposes 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 companies' CSR performance (Abriata and Delautre 2020, Jackson and Rathert 2016). The literature has also pointed at the importance of strong local public regulations and empowered civil society stakeholders for the effectivity of CSR commitments in global supply chains, especially for enabling rights, such as 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 (global, transnational, sectorial, national or company level) by actors of different natures (public or private) and with a variety of enforcement mechanisms (from purely voluntary to legally binding). This multiplicity brings the challenge of identifying the appropriate standards and methodologies to assess companies' practices on decent work as they operate in various environments. An interviewee (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 avoid discriminatory practices. Beyond legislation, cultural and other societal aspects complexify the comparative work. Several interviewees highlighted, for example, that the use of labour courts or strikes in labour disputes can be extremely variable from one country to another, making these indicators poor estimates of labour conditions in a comparative perspective. The continuously growing number of companies being rated is likely to make this issue even more critical.

The second challenge relates to objectivity. Most interviewees emphasised the need for these actors 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 criticised for their lack of

objectivity due to their reliance on analysts' interpretation (Chatterji and Levine 2006, Berg, Kölbel, and Rigobon 2020). Objectivity is an important criteria of professionalism and credibility (Stubbs and Rogers 2013). As IntRA10 said, "*we are not a research firm, we don't provide subjective opinion*". For that purpose, interviewees mentioned the need of providing "*neutral*" evaluations (IntE4) without "*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 suggest a level of objectivity.*" (IntRA8). Another important way to show objectivity is through audit trail, a system that traces data sources. The use of publicly available information was presented as essential to document decisions: "*Only publicly available data is used in our processes [...] because that is the only way we can show auditability and transparency*(IntRA10). At the same time, and as discussed earlier, public information tends to provide an overwhelmingly positive image of companies which spurs agencies to exchange with different stakeholders to have more reliable information, as underlined by an expert: "*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*" (IntA1). However, this kind of practices does not seem to be systematised: "*the intention was indeed to approach all stakeholders [...] but in reality, it was far less*" (IntRA6).

Yet, interviewees recognised that a certain level of subjectivity is inevitable when evaluating decent work. They regularly face ethical dilemmas such as where to set a threshold between good and bad practices, define scales regarding the severity of controversies or the diligence of companies in ensuring good working conditions: "*there is no limit to how much 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 comparison with the environmental dimension, ethical dilemmas seem to be more significant

when assessing decent work as it deals with human relations. Moreover, issues related to decent work appear for certain interviewees as politicised and sometimes holding the risk of favouring a "*Eurocentric, or Western-centric notion of what is good*" (IntRA6) because of the concentration of agencies in Europe and North America.

The nature of ESG rating agencies. The nature of agencies relates to their organisational form and the characteristics of the ESG data market. Two challenges were identified in this area. The first concerns the principle of materiality which, as explained earlier, brings the focus on what really matters in business terms and reduces the amount of unnecessary information. In the ESG context, the concept of materiality is theoretically expanded by considering a broader set of stakeholders (GRI 2013). These agencies' most important customers are financial actors-- asset managers, financial analysts, and investors—working under the logic of shareholder value. Therefore, assessing decent work is generally done "*from an investment perspective [...] what define the risks and opportunities for companies*" (IntRA2). Based on this logic, if an issue is not considered financially material it will be regarded as not relevant for investment. That does not necessarily mean that the issue is not relevant for decent work, but simply that there is 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). In the previous section, we saw how this tension can lead to very different technical arrangements between agencies concerning, for example, the inclusion of normative criteria in ESG items' weighting or in the rating of controversies. Similar to financial actors, the materiality principle tends to lock rating agencies in a short-term perspective, while the very definition of sustainability refers to the ability of organisations to balance the short and long terms. Some interviewees highlighted this

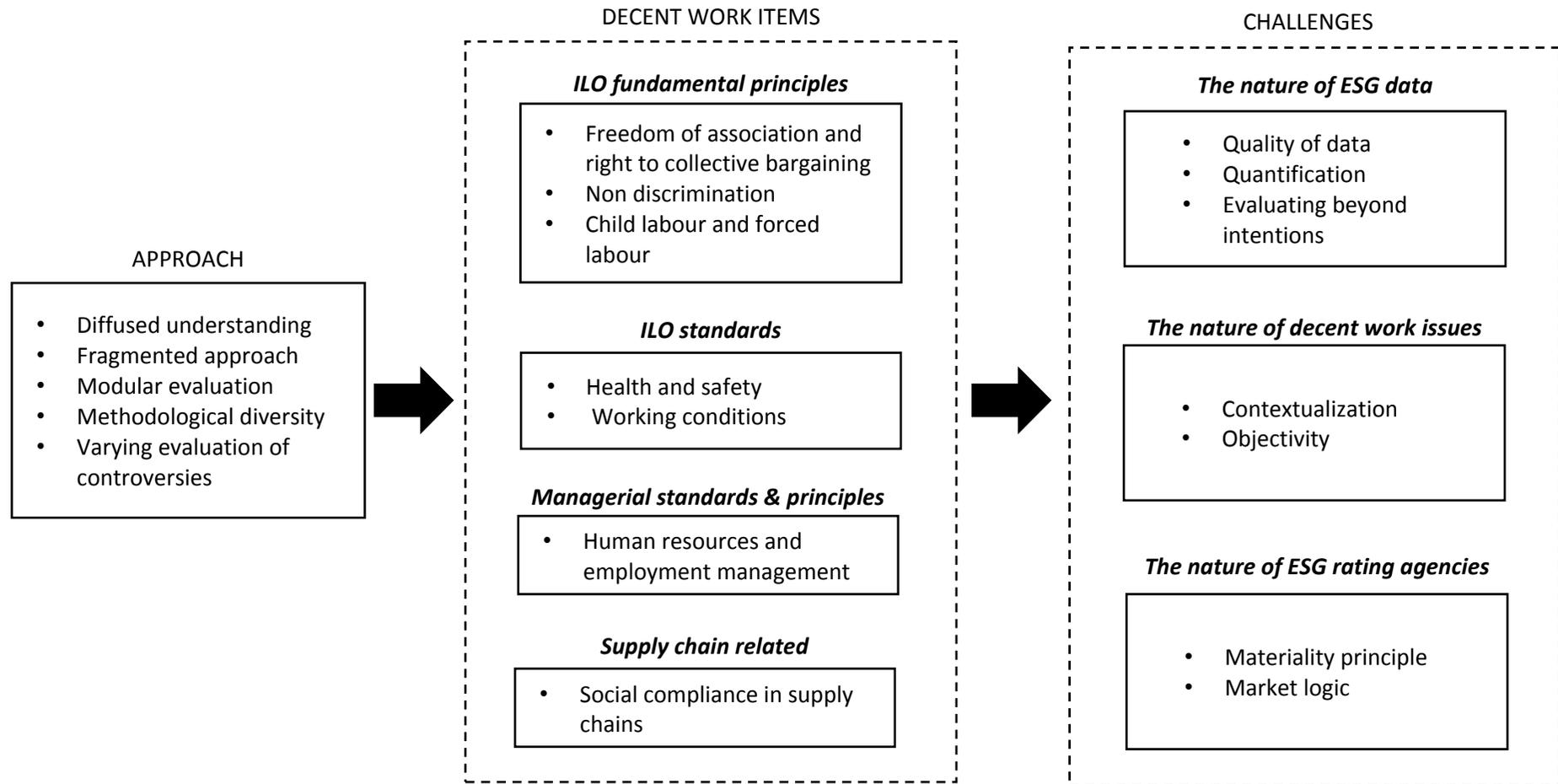
contradiction and its importance with regards to decent work: "*issues related to decent work might take more time to implement*" (IntE4).

The second challenge relates to the market logic. Rating agencies are for profit companies and face fierce competition. They must maintain a certain level of differentiation 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 be efficient to evaluate an ever increasing number of companies. As showed in Table 1, ESG rating agencies monitored between 4,700 to 11,000 firms. Yet, the lack of quantitative and clear data on decent work comes in opposition to this principle and makes the evaluation longer to carry out. An interviewee affirmed: "*Rating agencies do not have the capacity [in terms of time] to look deeply into these things.*"(IntRA1). Besides, rating agencies' customers have so far shown limited interest or mobilisation for decent work issues contrary to others like climate change: "*it is very seldom when they [the clients] demand concrete data on labour rights* " (IntRA4). Interviewees noted that more demand from clients would likely 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).

4. Discussion and Conclusions

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

Figure 2: Challenges to assess decent work



Our results show that decent work is a well-established and unquestionable dimension of ESG ratings. However, the notion of decent work itself remains diffused and abstract for many ESG professionals. If all agencies make explicit reference to ILO conventions and allocate a substantial share of their activity to measuring decent-work related items, none uses the term "decent work". They usually refer to different elements including freedom of association, child labour, gender equality, health and safety or working conditions rather than the general concept. As a result, those items are spread in multiple sections of the evaluation making it difficult to get a sense of companies' overall approach to decent work.

The 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 2020). This 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 2020, Scalet and Kelly 2010). We have also shown that there are differences regarding the scope, both in terms of breadth (the number of items covering decent work) and depth (the degree of details to evaluate each item). The methodologies for score aggregation, weighting, activation of items and controversies' assessment also diverge. Our findings also show that although rating agencies try to provide the most objective evaluations possible, some degrees of subjectivity are unavoidable but also desirable to assess decent work.

Evaluating and measuring decent work remains a cumbersome task for ESG professionals. A central problem is the lack of availability, reliability, and comparability of data (Chatterji et al. 2016, Christensen, Serafeim, and Sikochi 2020, Widyawati 2021). As previous studies showed, non-financial reporting differs significantly among regions and sectors (Fifka 2013, Ali, Frynas, and Mahmood 2017). Interviewees also highlighted the discrepancy between

policies and outcomes. Companies engage in what scholars identified as decoupling practices, through which they favour symbolic commitments over concrete measures (Bromley and Powell 2012, Tashman, Marano, and Kostova 2019). Recent research shows that increasing complexity of business models is a source of field opacity for lead companies and can generate more decoupling (Kuruvilla 2021, Wijen 2014). Other studies have shown that formulating policies may actually generate a sense of entitlement and a productivity narrative which may trigger action in companies (Haack, Schoeneborn, and Wickert 2012, Zeffane, Polonsky, and Medley 1994). It is therefore a significant starting point for companies to implement decent work policies. Still, overcoming the decoupling challenge is not easy. It could be done through site visits and in-depth and continuous dialogue with stakeholders. However, such approach would involve significant additional costs and time, which would affect agencies' business model. It also raises the question of ESG rating agencies' role. As an interviewee affirmed: "*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 in networks with grassroot organisations (Goodman et al. 2014) and experts in decent work could provide analysts with external views to triangulate the information provided by companies, and contribute in designing indicators to go beyond policy commitments. Although agencies work in partnerships with some organisations such as NGOs, collaborations remain limited. Collaborative platforms to share information and involve multiple actors, among which the rating agencies themselves, could be a powerful tool to assess and engage with companies on decent work-related issues.

Our study also shows 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 with the recent market consolidations. However,

quantitative data on decent work are scarce, which brings a serious limitation to what can be measured. Moreover, not all decent work issues are material in the financial sense or when considered in a short-term perspective. Such a strong focus and the preformatted rating grids could lead to a narrow outlook on decent work and lead to a loss of nuances in the evaluation. This relates to the importance of contextualising the information. However, quantified and material procedures show their limits when considering the complex regulatory framework in which companies operate at the transnational level.

In view of those challenges, one may wonder about the capacity of ESG rating agencies, and thereby RI, to address decent work even if this article does not call into question the very principle of rating. Agencies play an influential and important role in improving non-financial information disclosure. As an interviewees said, they are "*an accelerator for transparency*" (IntRA7). Notably, some investors-led initiatives focusing on decent work, such as the Workforce Disclosure Initiative, the Corporate Human Rights Benchmark, but also working groups within the PRI, work on improving disclosure of decent work related information and dialogue with companies. Moreover, by acquiring ESG rating agencies, mainstream credit agencies like Moody's, MSCI, S&P or 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 methodology and available data. They assess companies' intentions but can only partially evaluate policy outcomes. Controversies provide a necessary mechanism to counterbalance methodological limitations, but unfortunately companies are not equally targeted by the media and watch-groups. Additionally, if an event emerges as a controversy, it often means that serious damages have already happened.

Our study highlights the necessity for ESG data users to be actively engaged that is dialoguing with rating agencies, acting carefully when considering ESG data and understanding

ratings' logic and approach. However, there are several hindrances for such a proactive use of ESG information. First, agencies should be transparent on their underlying assumptions, methodologies, and data sources. As our study has shown, this is not always the case. In future years, it would be interesting to study the impact of the field consolidation on the quality and transparency of the ratings, especially for the social dimension, which is subject to more variety than other dimensions. Second, Amel-Zadeh and Serafeim (2018)'s 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 ESG issues per se. 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, it would be valuable to expand the scope of ESG information users to include, for example, trade unions or human rights organisations. Having a diversity of users would be beneficial to move beyond the financial logic of ESG information and require agencies to consider other dimensions in their evaluation.

Finally, it is also important to consider ESG rating agencies as part of a complex ecosystem where the actions of different actors are likely to mutually reinforce the promotion of decent work. The quality and reliability of ESG ratings depend largely on the ongoing movement encouraged by public and private actors to improve corporate reporting and transparency. If this movement weakens, this will have inevitably negative effects on the quality of ratings. Moreover, the capacity of ESG ratings (and more widely the capacity of the RI movement) to address decent work depends on the involvement of a series of actors including investors, companies and other stakeholders, including trade unions and other civil society groups. All these actors can push for a conceptualisation of materiality which better considers the complexities and the long-term perspective of decent work issues. Natural follow-ups to this research would be to analyse the inclusion of decent work issues in investors' strategies,

but also to 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 it. Some companies are doing pioneering work on decent work and are not necessarily well evaluated or recognised by rating agencies due to a mismatch between the evaluation criteria and their forward-looking approach.

More globally, and as a follow-up to the ILO Centenary Declaration 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 the structuring RI and ESG ratings market. In addition to their actions for an improved transparency of business actors on extra-financial issues, we could have also mentioned governments' role in regulating financial actors, especially when those actors manage funds coming from employees' saving for social protection purposes³. In addition to governments, social partners could also be considered for further research in the domain. As mentioned above, certain agencies can use unions as providers of extra-financial information to complement their sources. According to Penalva Icher (2008), unions can have a second role of representatives of employees' interests in the management of employee savings plans in many countries. Such a position could theoretically provide levers for unions in a strategy of shareholder activism. On their side, employers can be vocal critics of the absence of standardisation in the domain of ESG rating (Medef-Afep-Cliff-C3D 2019) at the national level. However, their interest to ensure a level playing field for how companies are assessed in their practices and policies regarding decent work should also encourage them to consider the issue at a more global level.

³ See for example, the ongoing debate in the USA concerning the [recent statement from the Department of Labor](#) on its rules on ESG investments for employees' retirement plans. See also Penalva-Icher (2008) on the case of France.

References

- Abriata, Bruno Dante, and Guillaume Delautre. 2020. What drives CSR? An empirical analysis on the labour dimensions of CSR. edited by ILO Working Paper 7: International Labour Organisation.
- Ali, Waris, Jędrzej 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. doi: 10.1002/csr.1410.
- 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. doi: 10.2469/faj.v74.n3.2.
- Avetisyan, Emma, and Kay Hockerts. 2017. "The consolidation of the ESG rating industry as an enactment of institutional retrogression." *Business Strategy & the Environment* 26 (3):316–330.
- Bansal, Pratima, and Mark R. DesJardine. 2014. "Business sustainability: It is about time." *Strategic Organization* 12 (1):70-78. doi: 10.1177/1476127013520265.
- Berg, Florian, Julian F. Kölbl, and Roberto Rigobon. 2020. "Aggregate Confusion: The Divergence of ESG Ratings." *Review of Finance* Forthcoming.
- Bourguignon, Rémi, and Arnaud Mias. 2017. Les accords-cadres internationaux: Etude comparative des ACI conclus par les entreprises françaises. edited by Report for the ILO Office for France. Geneva: ILO.
- 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'. Rochester, NY: Social Science Research Network.
- 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. doi: 10.5465/19416520.2012.684462.
- Brown, Barb, and Mike Wallace. 2018. The ESG Ecosystem Understanding the Dynamics of the Sustainability Ratings & Rankings Landscape. BrownFlynn.
- 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. doi: 10.1002/smj.2407.
- 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. doi: 10.1111/j.1530-9134.2009.00210.x.
- Chatterji, Aaron, and David Levine. 2006. "Breaking down the wall of codes: evaluating non-financial performance measurement." *California Management Review* 48 (2):29-51.
- Christensen, Dane, George Serafeim, and Anywhere Sikochi. 2020. Why is corporate virtue in the eye of the beholder? The case of ESG ratings. In *Working paper*, edited by Harvard Business School. Boston, MA.
- Corbin, Juliet, and Anselm Strauss. 2007. *Basics of qualitative research: Techniques and procedures for developing grounded theory* 3rd ed. Thousand Oaks, CA:: Sage.
- Creswell, John W. 1998. *Qualitative Inquiry and Design: Choosing Among Five Traditions*. London: Sage Publications.

- Creswell, John W. 2013. *Qualitative inquiry and research design: Choosing among five approaches* 3rd ed 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. doi: 10.5465/amp.2012.0123.
- DiMaggio, Paul J., and Walter W. Powell. 1983. "The iron cage revisited: institutional isomorphism and collective rationality in organizational field." *American Sociological Review* 48 (2):147-160.
- Dimmelmeier, Andreas 2020. Mergers and Acquisitions of ESG Firms: Towards a New Financial Infrastructure? : SocArXiv.
- 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. doi: 10.1057/jam.2015.31.
- Dumas, Christel, and Céline Louche. 2016. "Collective Beliefs on Responsible Investment." *Business & Society* 55 (3):427-457. doi: 10.1177/0007650315575327.
- Eccles, R. , and G. Serafeim. 2013. "The Performance Frontier: Innovating for a Sustainable Strategy." *Harvard Business Review* 91 (5):50–60.
- Eccles, Robert G., Linda-Eling Lee, and Judith C. Strohle. 2020. "The Social Origins of ESG? An Analysis of Innovest and KLD." *Organization & Environment* 33 (4):575-596.
- Eccles, Robert G., and Judith C. Strohle. 2018. *Exploring Social Origins in the Construction of ESG Measures*. Oxford, England: University of Oxford.
- Escrig-Olmedo, Elena , Maria Jesus Munoz-Torres, and Maria Angeles Fernandez-Izquierdo. 2010. "Socially responsible investing: sustainability indices, ESG rating and information provider agencies." *International Journal of Sustainable Economy* 2 (442-461).
- 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. doi: 10.1002/bse.729.
- Fortanier, Fabienne, Ans Kolk, and Jonatan Pinkse. 2011. "Harmonization in CSR Reporting." *Management International Review* 51 (5):665. doi: 10.1007/s11575-011-0089-9.
- Fowler, S., and C. Hope. 2007. "A critical review of sustainable business indices and their impacts." *Journal of Business Ethics* 73 (3):243-252.
- Gioia, D. A., K. G. Corley, and A. L. Hamilton. 2013. "Seeking qualitative rigor in inductive research: Notes on the Gioia methodology." *Organizational Research Methods* 16:15-31.
- Golafshani, N. 2003. "Understanding Reliability and Validity in Qualitative Research." *The Qualitative Report* 8 (4):597-606.
- Goodman, Jennifer, Céline Louche, Katinka Cranenburgh, and Daniel Arenas. 2014. "Social Shareholder Engagement: The Dynamics of Voice and Exit." *Journal of Business Ethics* 125 (2):193-210. doi: 10.1007/s10551-013-1890-0.
- Graafland, Johan, and Hugo Smid. 2019. "Decoupling Among CSR Policies, Programs, and Impacts: An Empirical Study." *Business & Society* 58 (2):231-267. doi: 10.1177/0007650316647951.
- GRI. 2013. *G4 Sustainability Reporting Guidelines*.: Global Reporting Initiative.
- GSIA. 2018. *2018 Global Sustainable Investment Review* Global Sustainable Investment Alliance.

- GSIA. 2021. 2020 Global Sustainable Investment Review Global Sustainable Investment Alliance.
- 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.
- IFAC. 2017. Making Sense of Materiality. International Federation of Accountants (IFAC).
- Ijiri, Y., and R.K. Jaedicke. 1966. "Reliability and objectivity of accounting measurements." *The Accounting Review* 41 (3):474-483.
- ILO. 1999. Decent Work. Report of the Director-General. International Labour Conference, 87th Session. . Geneva: International Labour Office.
- ILO. 2017. World Social Protection Report 2017-2019. Geneva: ILO.
- Jackson, Gregory, and Nikolas Rathert. 2016. "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*, 445-478. Emerald Publishing Limited.
- Köbel, Julian F., Florian Heeb, Falko Paetzold, and Timo Busch. 2019. Can Sustainable Investing Save the World? Reviewing the Impact of Investors on Companies. In *SSRN Scholarly Paper*. Rochester, NY: Social Science Research Network.
- Kuruvilla, S., Li, N., Lowell Jackson, J. . 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 G. Delautre, E. Echeverria Manrique and C. Fenwick. Geneva: ILO.
- 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." *ILR Review* 73 (4):841-872. doi: 10.1177/0019793920903278.
- 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. doi: 10.1007/s10551-018-3980-5.
- 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* 93:5-29.
- 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. doi: 10.1287/orsc.2015.1039.
- Medef-Afep-Cliff-C3D. 2019. Notation extra-financière : manifeste pour un meilleur dialogue avec les entreprises.
- Miles, M. B., A. M. Huberman, and J. Saldaña. 2014. *Qualitative data analysis: A methods sourcebook*. Thousand Oaks, CA: Sage.
- Nauman, Billy. 2019. "Credit rating agencies join battle for ESG supremacy." *Financial Times*, Moral Money. <https://www.ft.com/content/59f60306-d671-11e9-8367-807ebd53ab77>.
- Novethic. 2018. SDGs: a new focus for non-financial rating agencies. Novethic research centre.
- Penalva-Icher, Élise. 2016. "La Professionnalisation dans l'Investissement Socialement Responsable. Le cas des Analystes Extrafinanciers." *Revue Française de SocioÉconomie* 16 (1):141 - 159. doi: 10.3917/rfse.016.0141.
- 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. doi: 10.3917/rdli.057.0111.

- PRI. 2020. "An introduction to responsible investment ". Principles for Responsible Investment, accessed 05 December. <https://www.unpri.org/investment-tools/an-introduction-to-responsible-investment>.
- Scalet, Steven, and Thomas F. Kelly. 2010. "CSR Rating Agencies: What is Their Global Impact?" *Journal of Business Ethics* 94 (No. 1):69-88. doi: 10.1007/s10551-009-0250-6.
- Sekerler Richiardi, P. , and Ma. D. G. Arbo. 2019. Multinational enterprises and social protection: A case study of the L'Oréal Share & Care Program. edited by ILO Research Department Working Paper. Geneva: ILO.
- Semenova, Natalia, and Lars G. Hassel. 2015. "On the Validity of Environmental Performance Metrics." *Journal of Business Ethics* 132:249-258.
- Sjöström, Emma. 2008. "Shareholder Activism for Corporate Social Responsibility: What Do We Know?" *Sustainable Development* 16 (3):141-54.
- 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 C. 2019. "History of ESG Metrics." Papers on the State of ESG Standards, Yale, January 26.
- Stroehle, Judith Christina. 2017. "The enforcement of diverse labour standards through private governance." *Transfer: European Review of Labour and Research* 23 (4):475-493. doi: 10.1177/1024258917731016.
- Stubbs, W., and P. Rogers. 2013. "Lifting the veil on environment-socialgovernance 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. doi: 10.1057/s41267-018-0171-7.
- Widyawati, Luluk. 2021. "Measurement concerns and agreement of environmental social governance ratings." *Accounting & Finance* 61 (S1):1589-1623. doi: 10.1111/acfi.12638.
- 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. doi: 10.5465/amr.2012.0218.
- 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.