Environmental, Social and Governance evaluation for European Small and Medium Enterprises: A Multi-Criteria approach

October 4, 2024

Abstract

The exposure to environmental, social, and governance (ESG) risks can be effectively measured by companies to identify opportunities for long-term sustainable growth, along with the social and environmental impact. This process is crucial for listed small and medium-sized enterprises (SMEs) wanting additional support in their ESG transition, and for European SMEs it will be required by the implementation of the Corporate Sustainability Reporting Directive (CSRD), starting from 2026. In this contribution, we propose to apply a multi-criteria decision aiding approach to assess the sustainability profiles of SMEs. The methodology, which allows the measurements of a firm's ESG efforts (ESGness), is applied to a sample of European-listed SMEs, controlling for potential sector-specific effects, in order to understand what is the situation on the ESG front, and to identify ESG leaders and laggards. The model can provide valuable information for the firm, and for a broad spectrum of stakeholders, including policymakers and investors. The obtained rankings show some degree of robustness across different model parameterizations. The benefits of voluntary disclosure of sustainability information are investigated under a prudential scoring framework.

Keywords: Multiple criteria analysis, ESG, Small and medium-sized enterprises (SMEs), Sustainable policy

JEL Codes: C44, Q56, M14, O16

1 Introduction

The sustainability assessment, along with the Environmental, Social and Governance (ESG) dimensions, is becoming a critical issue for European SMEs that are subject to pressure from involved stakeholders (Torelli et al., 2020). Moreover, the Corporate Sustainability Reporting Directive (CSRD), that builds on the Non-Financial Reporting Directive (NFRD), shifts the focus from non-financial to sustainability reporting (Baumüller and Grbenic, 2021) including a broad set of new topics to be covered and it extends the scope of companies required to report on sustainability from 11,700 to approximately 50,000 (European Commission, 2023). The CSRD, which entered into force in January 2023, will be gradually applied to all listed EU companies, starting from reporting year 2026 for large companies to reporting year 2027 for listed SMEs, with the possibility to opt out of the reporting requirements for two more years.

More precisely, the new directive calls for a more comprehensive report on the impact of corporate activities on the environment and society, in line with the European Sustainability Reporting Standards (ESRS, 2023). Furthermore, under the CSRD a 'double materiality' assessment is mandatory, i.e. companies that report on sustainability need to assess the significance of sustainability issues from two angles. This approach requires to evaluate how the company's business and outlook is affected by sustainability risks (external perspective) and how the company's activities directly or indirectly impact the society and the environment (internal perspective), along each of the ESG dimensions. The environmental pillar addresses the commitment to reduce the environmental impact of a company's activity and in particular relies on saving and preserving natural energy and resources, assessing the carbon footprint and reducing the total greenhouse gas emissions, reducing total waste and preserving water scarcity. The social pillar addresses the relationships and interactions of a firm with all its stakeholders and, more broadly, with society. Among the issues tackled by this pillar we find employees, workforce health and safety, diversity and human rights. Finally, the governance pillar addresses the arrangement of rules and processes by which a firm is managed; among the key aspects, we mention the board diversity, the antibribery and corruption policies, training and education. Strong performance for the three ESG pillars, associated with a sound ESG profile, is having an increasing relevance and tangible effects on companies' overall financial health as documented by a growing body of research (see, for example, Höck et al. (2020), Friede et al. (2015)), making its assessment of paramount importance.

As Giese et al. (2019) show, in line with the seminal contribution of El Ghoul et al. (2011), ESG information can be reflected in a firm's valuation through lower cost of capital, higher profitability and lower systematic risk, under the assumption of no direct and indirect costs for ESG disclosure (Prencipe, 2004).

Understanding the firm's exposure to a range of potential ESG risks amounts to evaluating its *ESG performance*, i.e. the assessment of distinct ESG topics within a reporting framework such as the Global Reporting Initiative (2022), that can have a material impact on financial value creation for a company. Nonetheless, the modelization of the SMEs' ESG performance, despite its importance, has not received enough attention. Previous quantitative studies on SMEs have mainly investigated credit risk profiles (Voulgaris et al., 2000; Angilella and Mazzù, 2015; Corazza et al., 2016), whereas the literature on the role of ESG performance in SMEs has rather focused on the identification of drivers and challenges of the sustainability reporting initiatives (Ortiz-Martínez and Marín-Hernández, 2023). Other quantitative studies have a delimited sustainability problem in mind, such as waste management (Deshpande et al., 2020), but they do not propose a holistic quantitative framework and do not embrace the problem in general terms, for all the European SMEs.

To fill this gap, the research questions that guide our paper are the following:

- RQ1: What is the current situation on the ESG front for European listed SMEs, i.e. which firms are the leaders and laggards, relying on the outcomes of voluntary disclosure, in view of the CSRD application;
- RQ2: What is the role of undisclosed data in the ESG assessment problem and how such evaluation is affected by different assumptions about lack of disclosure;
- RQ3: What are the benefits of voluntary disclosure, and what are the implications of different disclosure decisions, from the point of view of the policymaker.

The first aim is to investigate the ESGness of SMEs, a task that is faced with several issues, in particular (i) the definition of a model, (ii) the construction of a dataset and, also w.r.t. RQ2, (iii) the management of undisclosed data. We deal with the first point by using a flexible partially-compensatory methodology that can account for a wide variety of goals, and that can provide valuable information to firms, markets and policymakers. Then, we also set up a unique dataset based on the identification of sustainability key performance indicators aligned with Global Reporting Initiative (2022), covering all the ESG dimensions

and all the listed European SMEs meeting minimum disclosure requirements. The collection of data is systematic and is constrained by the availability of indicators: in this way, an extensive analysis is performed to assess sectoral and geographical dimensions and to focus on the role of each ESG pillar. Finally, the robustness of the attained rankings is tested across different dimensions, namely w.r.t. to the significance of each pillar, to the DM's preferences and to the impact of undisclosed data. Such analysis is particularly timely and relevant, in light of the increasing commitment towards the sustainability dimensions of companies, the interest of the policymaker to govern transition to a new paradigm, and to steer the firms' attitude towards sustainability accordingly, given also the importance for both actors to convey ESG information to financial markets for a more comprehensive evaluation. The relevance of the sustainability assessment has been highlighted in early attempts by Ortiz-Martínez and Marín-Hernández (2022, 2023): to characterize the progress on sustainability topics of SMEs they have focused on lexical properties of reports, the adherence level to GRI standards or the presence of external assurance. The second aim, strictly related to the issue of limited and inconsistent reporting of data in SMEs (Angilella and Mazzù, 2015), is to evaluate the adoption of a suitable prudential rule for imputation and the assessment of the scoring procedure's robustness to different imputation techniques. Finally, the third aim is to gauge the benefits of voluntary disclosure, by evaluating two scenarios faced by a company willing to take an optimal disclosure decision, assuming either that all the other firms fully disclose ESG information, or that there is at least another firm that does not report its performance on a specific topic.

In practical terms, a systematic assessment of the sustainability profiles of SMEs is crucial as it has several implications for stakeholders and the firm itself, since (i) it involves disclosing price-sensitive information to investors, (ii) provides the policymaker with additional tools for governing the transition to sustainable processes, and (iii) equips firms with a relevant transition/regulatory risk metric. However, the construction of an ESG score for SMEs holds significant relevance also within the framework of agency and stakeholder theories (Freeman and Reed, 1983; Bofinger et al., 2022; De Falco et al., 2024), as it enhances transparency, by addressing all the stakeholders and mitigating conflicts between principals and agents; hence, our proposal can be thought as a quantifiable measure of how well a company aligns with societal expectations, eventually gaining or maintaining legitimacy in the eyes of stakeholders, and ultimately alleviating related environmental, social or reputational risk (Singhania and Gupta, 2023).

The rest of the paper is organized as follows. Section 2 briefly presents research trends

in the literature on financial risk management and sustainability in SMEs. Section 3 presents how data has been collected and pre-processed, whereas Section 4 presents the methodology. Section 5 is devoted to the presentation and discussion of the results, while Section 6 discusses some possible implications of our findings. Finally, Section 7 concludes and presents some future research directions.

2 Literature review

The peculiar features of SMEs have drawn the interest of scholars and practitioners alike, especially with respect to the creditworthiness assessment problem (Altman and Sabato, 2007), since the credit supply is among the primary transmission channels of economic shocks for SMEs (D'Amato, 2020). The inclusion of non-financial information is potentially as crucial as standard financial ratios used for credit risk assessment when assessing credit or equity financing for such companies.

A strand of the literature has highlighted the necessity of developing specific credit risk models for innovative SMEs (Czarnitzki and Hottenrott, 2011), including a range of econometric and Multiple Criteria Decision Aid (MCDA) models. In this case, lending relies necessarily on soft information (Moro and Fink, 2013).

Previous studies have emphasized the size of a company itself as one of the key drivers for the adoption of corporate social responsibility (CSR) standards (Goyal et al., 2013).

As far as the benefits are involved, compared to large corporations, the managerial structure of SMEs is simpler: agency costs are therefore smaller or missing, potentially because the agent and the principal coincide or direct supervision occurs (Bartolacci et al., 2020). From the perspective of costs, the picture is less clear. As Gjergji et al. (2021) argue, direct and indirect economic costs are among the main barriers to the development of CSR in SMEs. Companies indeed bear direct costs due to required expertise and investments in sustainability reporting (production and dissemination costs) and indirect costs from disclosing segment information (competitive costs). Both are larger for SMEs (Prencipe, 2004), mainly because of the larger fixed component of costs and the difficulty for small firms in protecting from competitors. On the contrary, Rodríguez-Gutiérrez et al. (2021) claim that costs play a marginal role in the ESG disclosure decision, suggesting that there is widespread awareness among SMEs that CSR can yield long-term returns and that economic cost is a required precondition.

Nonetheless, the preparation of a sustainability report requires a decision-making pro-

cess characterized by standard rules: lack of operational tools, along with technical knowledge and advertising skills (Gjergji et al., 2021) are major concerns for all SMEs that wish to formalize such sustainability practices.

To summarize, if the impact of ESG disclosure is widely recognized as positive for large firms' valuation (El Ghoul et al., 2011), since voluntary non-financial information can effectively mitigate exposure to a broad range of direct and indirect risks, the relationship is way less clear for SMEs.

Due to SMEs' scant track records in sustainability reporting, the MCDA approach (Zionts, 1979) is particularly suitable for assessing the ESG performance of firms, thanks to its flexibility. For credit rating models, a large volume of studies is available for both traditional and innovative SMEs (Voulgaris et al., 2000; Angilella and Mazzù, 2015; Corazza et al., 2015, 2016).

The governance performance of large-cap companies is investigated by Guney et al. (2020); they construct a competing corporate governance quality indicator through a well-known MCDA approach, and find that it is robust across different subsamples, corporate performance indicators, and industries.

3 Data analysis

In what follows, we illustrate in detail the construction and cleaning process of the hand-collected dataset used in the analysis.

Our reference source for identifying the perimeter of European listed SMEs is the Orbis database of Bureau van Dijk. In accordance with the definition established by the European Commission, we consider active firms with less than 50 million turnover and less than 250 employees. The hand-collected dataset we obtain is composed of 1,337 listed European SMEs, retrieved on 26 April 2023. Practically speaking, this implies that most firms included in the sample are, according to standard practice and from a market capitalization point of view, microcap companies. As a result, such companies are characterized by limited coverage from analysts and diverse rating providers, implying that little to no commonality can be found in the definition of ESG attributes and standards. Hence, the ESG performance assessment must necessarily rely on a combination of human supervision and to a certain extent, also of systematic retrieval of data, otherwise no assessment of firms' ESG performance could be made.

The screening process discussed in detail below requires adjustments of the reported

number, so as to guarantee homogeneous comparisons among firms. Companies reporting data for more than 50% of selected criteria are kept. If the disclosed data of a firm cannot be harmonized, in accordance with the definitions provided in Table 2, then the value is reported as missing. Out of $1,337$ firms of 27 European countries belonging to 11 sectors, we obtain a unique hand-collected dataset with 811 observations (104 firms observed across 12 different criteria, with around 35% missing data) for a cross-sectional study. The countries and the sectors for which at least one firm releases sustainability information are reported in Table 4.

Determining sustainability indicators that are informed by different voluntary reporting standards and frameworks across sectors and countries is a complex task, which can be boiled down to (i) assessing the double materiality of each topic and (ii) slimming down the list of candidate topics in order to establish a level playing field for all companies.

Despite the importance attributed to materiality assessment in the literature, due to limited willingness to disclose CSR information caused by a poor cost-benefit ratio of CSR reporting for SMEs, a thorough evaluation of relevant criteria becomes quite complicated (Gjergji et al., 2021).

The choice of criteria is based on an overall evaluation of all the reports according to the Global Reporting Initiative (2022) standards. A non-discretionary assessment of the relevant ones is necessary: thus, a thorough screening of the sustainability reports from the initial sample of listed European SMEs has been performed, then possible candidates have been filtered by removing all the criteria characterized by a boolean value (e.g. having a whistleblowing system in place), since they were found to provide little added value to the analysis. The chosen criteria have been applied to the entire SMEs sample, irrespectively of the country and sector. From this set of companies, we have collected all the relevant sustainability reports, which are available for around the 20% of the initial set of firms. More precisely, data are screened by selecting first all the relevant observations, according to a predefined set of keywords reported in Table 1, partly derived from the publicly-available dictionary of Baier et al. (2020).

The description, the optimization goals and the purpose of criteria is reported in Table 2, along with a criteria-GRI principles mapping, where the environmental, social and governance topics are respectively denoted as E_{ij}, S_{ij} and G_{ij}, where i, j identify respectively the criterion and the firm.

Moreover, in Table 3 we present the formal construction of criteria, previously introduced in Table 2. For three out of four environmental criteria, we normalize by company

Table 1: List of keywords for retrieving j observations from sustainability reports

Criterion	Keywords
E_{1j}: Carbon intensity	*'GHG'*, *'tCO_2e'*, *'Carbon emissions'*, *'Scope 1 and 2'*
E_{2j}: Waste generation intensity	*'Waste management'*, *'Hazardous waste'*, *'Waste management'*
E_{3j}: Renewable electricity consumption	*'Electricity consumption'*,*'Renewable electricity '*, *'Electricity production'*, *'kWh'*
E_{4j}: Water consumption intensity 4	*'Water consumption'*, *'Cubic meter'*, *'m^3'*
S_{1j}: Average training hours	*'Training hours'*, *'Learning'*, *'Training course'*
S_{2j}: Job creation	*'Employee turnover'*, *'Recruitements'*, *'Terminations'*, *'Job creation'*
S_{3j}: Management diversity by gender	*'Female managers'*, *'Female middle-managers'*, *'Ratio of female to male managers'*
S_{4j}: Gender pay gap	*'Gender pay gap'*, *'Salary inequality'*, *'Gender wage gap'*, *'Gender wage ratio'*
G_{1j}: Board diversity by gender	*'Board diversity'*, *'Percentage of female in boards'*, *'Board composition'*
G_{2j}: Economic value generation and distribution	*'Value distribution'*, *'Economic value distributed'*
G_{3j}:Board independence	*'Independent board directors'*, *'Board independence'*
G_{4j}: CEO pay ratio	*'CEO pay gap'*, *'CEO to average wage ratio'*

revenues in million euros. The third criterion, E_{3j} is equal to the percentage of non-renewable electricity consumption. As for the social criteria, we divide by the number of employees (see e.g. criteria S_{1j} and S_{2j}) or we consider the absolute value of a distance to assess gender parity. The definition of governance standardization criteria is also straightforward, since we consider the absolute distance from gender equality, as well as revenues and the number of board members. The average salary is used as a standardizing criterion to compute the CEO pay ratio.

As far as the disclosure process is concerned, note that the breakdown of the sample reported in Table 4 by sector and country is characterized by significant concentration, with two countries making up exactly the 50% of the overall sample size. Note also that after performing a NACE-to-GICS mapping of sectors, two are not included in our sample, i.e. Energy and Materials: being highly capital-intensive sectors, the reason for this is intuitive and the result unsurprising. Despite considering a small sample, our findings w.r.t. the composition of the sample are aligned with the literature, and could be explained in various ways. The standards of ESG reporting differ significantly across countries, industries, and firms. Moreover, it has been shown that peer effects and the stakeholders' commitment to ESG targets and CSR strategies carried out by competitors in a given ecosystem are known for impacting on the quality of ESG reporting (Torelli et al., 2020).

Altogether, we document that limited disclosure is available w.r.t. the employees' pay, in particular when taking into account the gap between female and male employees, as well as with their CEOs. Furthermore, we find that the number of missing values is approximately stable across countries.

Table 2: List of criteria. The cardinality, the optimization goals, the correspondance with the GRI mapping and a concise description are respectively reported for each criterion, for the alternative j.

Criterion	Cardinality (#)	Goal	GRI	Description and purpose
E_{1j}: Carbon intensity	70	Min	305-4	Carbon intensity captures both direct emissions from owned or controlled sources (scope 1) and indirect emissions from the generation of purchased electricity (scope 2). It corresponds to the ratio of Greenhouse Gases (tCO_2e) emissions to the firms' revenues in €. Emissions indirectly generated by the company's value chain (scope 3) are not considered due to lack of reliable estimates. The exposure to climate risk is captured.
E_{2j}: Waste generation intensity	44	Min	306-1	The waste generation intensity is defined as the ratio of hazardous and non-hazardous waste generation to the firm's revenues.
E_{3j}: Non-renewable electricity consumption (%)	95	Min	302-1	The criterion is computed as the ratio of non-renewable electricity to total electricity consumption (in KWh), allowing to assess commitment to an environmentally sustainable supply chain.
E_{4j}: Water consumption intensity	48	Min	303-5	The water generation intensity is defined as the ratio of water consumption to firm's revenues.
S_{1j}: Average training hours	40	Max	404-1	The criterion corresponds to the ratio of overall training hours (vocational training, instruction and training or education pursued externally) to the overall headcount (middle and top management included). Commitment to high quality working conditions is captured.
S_{2j}: Job creation (%)	94	Max	401-1	The criterion corresponds to the ratio of the difference between new hires and terminations at time $t+1$ to the overall headcount at time t (interns included).
S_{3j}: Management diversity by gender (%)	83	Min	405-1	The criterion is defined as the ratio of female managers to the total number of managers and the gender equality level (i.e. 50%).
S_{4j}: Gender pay gap	20	Min	405-2	The gender pay gap is defined as the absolute difference between average gross annual wage of male and female employees, without controlling for seniority, so as to evaluate the firm's commitment to gender parity.
G_{1j}: Board diversity by gender (%)	104	Min	405-1	The criterion is defined as the ratio of female directors to the total number of directors and the gender equality level (i.e. 50%).
G_{2j}: Economic value generation and distribution (%)	72	Max	201-1	The criterion is defined as the ratio of economic value distributed by the company to its stakeholders (suppliers, employees, lenders, public administration and shareholders) to the economic value generated, i.e. the annual turnover. This value represents the wealth produced by the firm and its impact on key stakeholders.
G_{3j}: Board independence (%)	104	Max	102-22	The criterion is defined as the ratio of independent directors to the overall number of directors.
G_{4j}: CEO pay ratio (%)	37	Min	102-38	The criterion corresponds to the ratio of the annual remuneration paid to the CEO to the average annual remuneration of all employees.

Table 3: Measurement of criteria, with j denoting alternatives.

Criterion	Measurement
E_{1j}: Carbon intensity	$\mathcal{CI}_j = \frac{\mathcal{CE}_j}{\mathcal{R}_j}$
E_{2j}: Waste generation intensity	$\mathcal{WGI}_j = \frac{\mathcal{WG}_j}{\mathcal{R}_j}$
E_{3j}: Renewable electricity consumption (%)	$\mathcal{RS}_j = \frac{\mathcal{RE}_j}{\mathcal{E}_j}$
E_{4j}: Water consumption intensity	$\mathcal{WCI}_j = \frac{\mathcal{WC}_j}{\mathcal{R}_j}$
S_{1j}: Average training hours	$\overline{\mathcal{T}}_j = \frac{\mathcal{T}_j}{\mathcal{EM}_j}$
S_{2j}: Job creation (%)	$\mathcal{JC}_j = \frac{\mathcal{EM}_{t+1,j} - \mathcal{EM}_{t,j}}{\mathcal{EM}_{t,j}}$
S_{3j}: Management diversity by gender (%)	$\mathcal{MD}_j = \lvert\frac{\mathcal{FM}_j}{M_j} - 0.5\rvert$
S_{4j}: Gender pay gap	$\mathcal{GG}_j = \lvert\mathcal{MR}_j - \mathcal{FR}_j\rvert$
G_{1j}: Board diversity by gender (%)	$\mathcal{BD}_j = \lvert\frac{\mathcal{FB}_j}{\mathcal{MB}_j} - 0.5\rvert$
G_{2j}: Economic value generation and distribution (%)	$\mathcal{EVD}_j = \frac{\mathcal{RD}_j}{\mathcal{R}_j}$
G_{3j}: Board independence (%)	$\mathcal{BI}_j = \frac{\mathcal{IBM}_j}{\mathcal{BM}_j}$
G_{4j}: CEO pay ratio (%)	$\mathcal{R}_{\mathcal{CEO}/\mathcal{EM},j} = \frac{\mathcal{R}_{\mathcal{CEO},j}}{\overline{\mathcal{R}}_{\mathcal{EM},j}}$

As for the remaining topics, we find that a relatively high percentage of companies disclose sustainability-related information, although some criteria, such as the number of independent directors, are also affected by mandatory financial requirements. Based on a thorough screening of reports, we also find that the number of missing values for other criteria, such as the number of training hours per employee, are attributable to peer-effects, country, or ecosystem-specific reporting standards (Torelli et al., 2020). For instance, note that w.r.t. the training hours, more than 50% of the available data are referred to Italian companies, while more than two-thirds of firms reporting on the CEO pay ratio are Swedish. In addition, aside from mandatory requirements, we find that companies tend to limit disclosure to a restricted scope in relation to governance topics. The resulting difficulty in extracting drivers for assessing the governance performance could also be related to the negative relationship between firm performance and corporate governance quality (Guney et al., 2020) found in the literature, especially with respect to board diversity in SMEs (Shehata et al., 2017).

Table 4: Sample size of companies disclosing ESG information by country and sector. Note that EU countries without relevant data are not included in the table; in parenthesis, the overall number of listed SMEs in the EU is reported. The breakdown of firms by sector is the result of a NACE-to-GICS mapping performed by the authors.

Country \ Sector	Industrials	Consumer discretionary	Consumer staples	Health Care	Financials	Information Technology	Communication services	Utilities	Real Estate	**Total**
Austria				1						1 (4)
Belgium	1								1	2 (19)
Denmark	4	1			5	2			1	13 (59)
Estonia									1	1 (8)
Finland	2			1	1	1				5 (37)
France	5			5	1	1		1	1	14 (191)
Germany		1	1			1	1			4 (120)
Hungary					1			1		2 (15)
Italy	3	5		3	2	8	3	1	1	26 (226)
Latvia		1								1 (5)
Netherlands			2							2 (18)
Poland									3	3 (211)
Romania	1									1 (97)
Spain								2	1	3 (46)
Sweden	10	2		3	2	3			6	26 (220)
Total	26	10	3	13	12	16	5	5	15	**104** (**1,337**)

4 A multicriteria decision model for ESG assessment

In light of the lack of quantitative approaches for modeling ESG profiles for SMEs, in this section, we propose to cope with such problem by using an MCDA approach (Zionts, 1979). MCDA methods allow us to deal with unstructured decision making, in a flexible way by assessing the impact of multiple conflicting criteria. More precisely, in what follows, we resort to the MURAME model (Goletsis et al., 2003), which allows us to make a very limited number of assumptions with regard to the involved parameters and missing data. In Section 4.1, we illustrate the methodology used in this work, a multicriteria outranking methodology (MURAME) which merges two deep-rooted multi-criteria decision-aiding methodologies, namely ELECTREE III (Figueira et al., 2016) and PROMETHEE II (Brans

and Vincke, 1985). Finally, in Section 4.2, we relate the characteristics of the data to the model.

4.1 A multicriteria ranking method (MURAME)

Let us consider the preference structure of a decision-maker (DM). Given a set of m alternatives $\mathcal{A} = \{a_1, \ldots, a_i, \ldots, a_m\}$ and a set of n criteria $\mathcal{C} = \{c_1, \ldots, c_j, \ldots, c_n\}$, the alternatives are assessed according to a decision matrix $\mathbf{G}_{m \times n}$, whose elements g_{ij} indicate the performance for the alternative i, w.r.t. the criterion j.

The DM's uncertain preferences are taken into account by means of an indifference threshold, q_j, a preference threshold p_j and a veto threshold v_j, for any given criterion c_j; in any case, it holds that $q_j \leq p_j \leq v_j$. Let us begin by defining, for each pair of alternatives, the local concordance $C_j(a_i, a_k) \in (0, 1)$ and discordance $D_j(a_i, a_k) \in (0, 1)$ indexes in Equations (1)-(2), as follows:

$$C_j(a_i, a_k) = \begin{cases} 1 & g_{kj} - g_{ij} \leq q_j \\ 0 & g_{kj} - g_{ij} \geq p_j \\ \dfrac{g_{ij} - g_{kj} + p_j}{p_j - q_j} & \text{otherwise} \end{cases} \tag{1}$$

$$D_j(a_i, a_k) = \begin{cases} 0 & g_{kj} - g_{ij} \leq p_j \\ 1 & g_{kj} - g_{ij} \geq v_j \\ \dfrac{g_{kj} - g_{ij} - p_j}{v_j - p_j} & \text{otherwise} \end{cases} \tag{2}$$

The local concordance index quantifies by how much the alternative a_i is at least as good as the alternative a_k, whereas the discordance index is used to gauge the extent to which a_k dominates a_i. If $g_{kj} - g_{ij} \geq v_j$, then the DM expresses a strict preference for the alternative a_k over the alternative a_i.

The global concordance index in Equation (3) is computed through the aggregation of the local concordance indexes, as follows:

$$C(a_i, a_k) = \sum_{j=1}^{n} w_j C_j(a_i, a_k), \tag{3}$$

where w_j represents the normalized non-negative weight associated with a given criterion

j.

An outranking index $O(a_i, a_k)$ is constructed in Equation (4) by putting together information from both discordance and global concordance indexes:

$$O(a_i, a_k) = \begin{cases} C(a_i, a_k) & D_j(a_i, a_k) \leq C(a_i, a_k) \quad \forall j \\ C(a_i, a_k) \prod_{j \in \mathcal{J}} \frac{1 - D_j(a_i, a_k)}{1 - C(a_i, a_k)} & \text{otherwise,} \end{cases} \tag{4}$$

where $\mathcal{J} \subseteq \{1, \ldots, n\}$ is the subset of criteria such that $D_j(a_i, a_k) > C(a_i, a_k)$.

Finally, a final ranking of alternatives is recovered by computing a net flow $\varphi(a_i)$, which amounts to taking the difference between out and inflows, denoted with $\varphi^+(a_i)$ and $\varphi^-(a_i)$, for each alternative a_i. Such quantities allow to measure respectively the strength and the weakness of an alternative a_i relative to the others:

$$\varphi(a_i) = \varphi^+(a_i) - \varphi^-(a_i) = \sum_{k \neq i} O(a_i, a_k) - \sum_{k \neq i} O(a_k, a_i). \tag{5}$$

Hence, the alternatives can be ranked according to the net flow φ and normalized so as to attach a score $\mathcal{S} \in [0, 100]$ to each alternative.

4.2 Relating the model to the data

The MURAME is a parsimonious model requiring a limited number of assumptions w.r.t. the DM's preferences. However, the characteristics of the data and the imputation method can still have a notable impact on intermediate computations.

First of all, let us consider the issue of missing data, whose impact has been extensively studied by a large volume of both theoretical (Tsikriktsis, 2005) and ESG-related contributions (Kotsantonis and Serafeim, 2019; Sahin et al., 2022). Note that the measurement of ESG ratings is quite ambiguous, and existing definitions are sometimes competing and unclear; there is little to no agreement on the true sustainability drivers (Billio et al., 2021), which are moreover plagued by missing data at the company level.

In light of the specificities of the outranking model we use, we attribute the worst sector value to companies that do not release data, and if there are no available data in a given sector, we impute the worst global value. We want to avoid penalizing firms that already disclose data, due to firms that underreport ESG information, which would occur, when performing pairwise comparisons between firms over different criteria, if softer imputation

approaches were adopted. Furthermore, note that the missing information could actually be a negative signal about the firm, hence it might be safer to assume that it is doing poorly (Lindsey et al., 2022).

As we discuss in Section 5.2.2, it might not be convenient to disclose sustainability information voluntarily. Therefore, it is not completely unreasonable to assume that if the model design is publicly available, then poorly performing firms w.r.t. a given criterion, might not be encouraged to report such information, resulting in a self-selection bias. From a normative point of view, this provides a case against attributing an additional penalty for missing observations, which is also the same line of reasoning of the approach discussed in (Sahin et al., 2022), whereas for alternative imputation approaches we refer to Kotsantonis and Serafeim (2019).

Moreover, in order to make a fair comparison among firms, we adopt a 'best-in-class' ESG approach[1] (Henriksson et al., 2019), hence data are normalized by sector in a range $[0, 1]$, so as to compensate for the impact of sector-specific features. Better coverage and higher data quality would allow to explicitly incorporate sector and country effects into the model, in line also with recent contributions in the field of corporate default risk prediction (Doumpos et al., 2017). Therefore, we perform a linear transformation on the original data, with j denoting the alternatives for a given criterion and $\theta \in \{$*Industrials, Consumer discretionary, Consumer staples, Health Care, Financials, Information Technology, Communication services, Utilities, Real Estate*$\}$ corresponds to the sector the alternative j belongs to:

$$X_{scaled,\theta j} = \frac{X_{\theta j} - X_{min,\theta j}}{X_{max,\theta j} - X_{min,\theta j}} \tag{6}$$

In case there are no observations available for a specific sector w.r.t. a criterion, then the global maximum and minimum value of the entire sample are used for normalization.

Let us now briefly present how the methodological assumptions are directly related to our dataset. In Table 5, we report the unconditional correlations between criteria: note that in most cases the correlations are close to zero, with a few exceptions.As far as the correlation between criteria is concerned, although multicriteria decision aiding does not suffer from multi-collinearity issues, we report an ex-post assessment of Pearson correlations since a large number of variables might undermine the interpretability of the model, especially if they are highly correlated (Corazza et al., 2016).

[1] A best-in-class approach in sustainable investing amounts to screening companies by finding leaders in their sector or industry.

Table 5: Unconditional correlations between criteria.

	E_1	E_2	E_3	E_4	S_1	S_2	S_3	S_4	G_1	G_2	G_3	G_4
E_1	1.0000	0.4739	0.0215	0.3259	0.1288	0.0689	−0.0644	−0.2253	0.0035	−0.1215	−0.1024	0.0771
E_2		1.0000	−0.0321	0.2951	−0.0014	0.1889	−0.0218	−0.3088	0.1560	−0.0380	−0.1416	−0.0358
E_3			1.0000	0.1542	−0.0511	−0.2127	0.0648	0.0064	0.1676	−0.0949	−0.1801	−0.0195
E_4				1.0000	−0.0554	0.0188	−0.0265	−0.2963	−0.0075	−0.1392	0.0476	−0.0819
S_1					1.0000	0.0775	−0.0361	−0.0931	0.0657	−0.0723	−0.1595	0.0733
S_2						1.0000	0.1462	0.0061	0.1184	0.1431	−0.0415	0.1296
S_3							1.0000	0.2567	0.1884	−0.0639	−0.1409	0.2408
S_4								1.0000	−0.0894	−0.0693	0.0619	0.3580
G_1									1.0000	0.0227	−0.1666	0.0987
G_2										1.0000	0.0824	−0.0041
G_3											1.0000	0.1137
G_4												1.0000

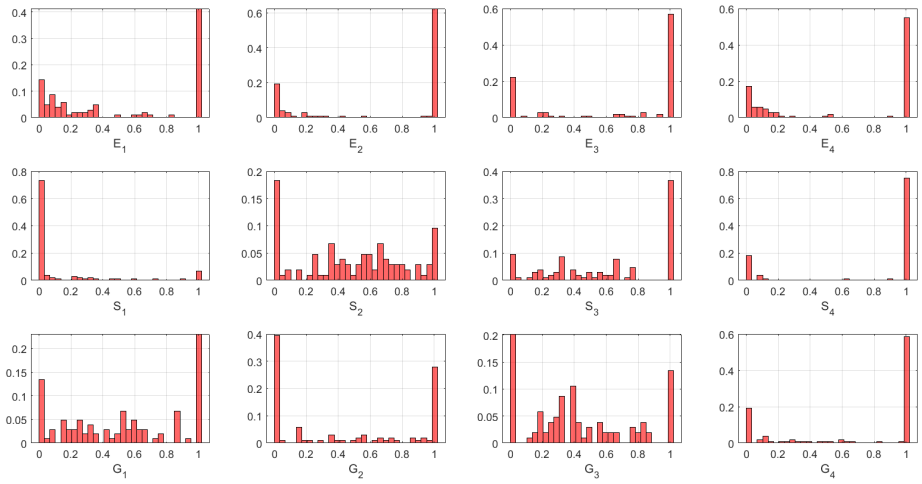

Figure 1: Performance distributions of criteria estimate via histograms w.r.t. normalized and imputed datasets.

Finally, in Figure 1 we report the performance distributions of criteria: note that the imputation procedure moderately affects the shape of the distribution of specific criteria, either on the right or the left of the distribution, respectively, depending on whether the criteria is minimized (E_1, E_2, E_3, E_4, S_3, S_4, G_1, G_4) or maximized (S_1, S_2, G_2, G_3). Further considerations about the impact of the imputation choices on the scoring procedure are left to Section 5.

5 Application to the ESG scoring problem of SMEs

In what follows, we apply the scoring procedure to the dataset of European SMEs. First, in Section 5.1 we discuss the scores derived from the procedure, assuming that each aspect of ESG profiles assessment is equally important after sector-specific normalization; then, in Section 5.2 we perform a sensitivity analysis to evaluate the robustness of the results. We conclude by briefly discussing the policy implications of disclosure choices in Section 5.3.

5.1 Scoring procedure and breakdown by pillar

In this section, we discuss the rating of firms across different sustainability dimensions, including both the overall standings and the score within each pillar.

As for the parameter settings, we keep subjectivity at a minimum. Equal weights are applied to each criterion, and the thresholds q_j, p_j, and v_j are set respectively equal to the first, the third, and the fourth quintile of the distribution of criterion j. Such choice of parameters is further supported by a sensitivity analysis presented in Section 5.2.

In Table 6, we present the results with regard to the top ten and worst ten performers in the column ESG. We use a palette of ten colors to identify exactly ten ranges of normalized scores $\mathcal{S} \in [0, 100]$, from $[0, 10)$ to $[90, 100]$. It is worth noting that, by construction, there is not a linear relationship between aggregate ESG scores $\varphi_{net,ESG}$ and the scores for each pillar $\varphi_{net,i}$, with $i \in \{E, S, G\}$ (see Equations (3)-(4)), reported respectively in the columns E, S and G. Indeed, the scores of the E, S, G columns are obtained by running the model three times separately for each pillar i, where we set the weight for pillar i (i.e. a subset of criteria belonging to i) equal to 1. Thus, by attributing a weight only to a single pillar i, the sustainability performance of firms in columns E, S, G should be interpreted as if the DM attached importance to only one of the three profiles.

Table 6: Top and bottom ten firms according to the ESG scoring procedure. The overall score is reported along with a breakdown of scores by pillar. Note that, by construction, the overall score based on the net flow φ_{net} is non-compensatory.

Overall Ranking	ID	Sector	Country	ESG	E	S	G
1	57	Utilities	Spain	100.00	100.00	100.00	93.79
2	91	Industrials	Sweden	77.92	100.00	75.71	77.21
3	42	Industrials	Finland	76.12	86.94	78.78	78.09
4	11	Communication Services	Italy	71.88	72.42	89.73	76.06
5	58	Utilities	Spain	67.45	80.61	78.94	74.53
6	70	Industrials	France	66.81	81.12	75.23	66.34
7	9	Communication Services	Italy	66.11	80.07	82.38	73.04
8	93	Industrials	Sweden	65.32	65.07	61.08	88.28
9	56	Communication Services	Spain	64.59	66.20	66.06	88.95
10	8	Consumer Discretionary	Italy	64.16	99.59	83.39	37.01
95	68	Information Technology	France	22.51	35.52	40.54	32.67
96	50	Industrials	Netherlands	21.21	0.87	37.29	38.17
97	13	Real Estate	Italy	20.88	0.75	16.56	37.93
98	38	Information Technology	Denmark	17.11	31.19	31.31	24.15
99	16	Financials	Italy	13.36	0.31	52.89	18.02
100	26	Health Care	Italy	10.58	0.85	39.83	20.62
101	23	Information Technology	Italy	1.93	0.04	18.61	14.20
102	75	Industrials	Romania	1.15	16.08	4.74	15.97
103	53	Consumer Discretionary	Germany	0.91	0.00	18.58	0.00
104	28	Financials	Denmark	0.00	14.81	15.77	15.44

As for the results in Table 6, note that the first firm outperforms the others by a wide margin, also thanks to a strong performance across all dimensions, since it is ranked in the top five also within single sustainability dimensions. Among the leading firms, the utilities and communication services sectors stand out. As for the former, a preliminary analysis of all the collected sustainability reports shows that renewable energy companies are more prone to release ESG data. With respect to geographical and sectoral clusters, note also that all the Spanish firms included in our sample are ranked in the top ten, whereas among the laggard firms, note that IT SMEs, despite releasing more information compared to firms in other sectors, tend to perform poorly on average. Instead, utility firms are well represented at the upper end of the rankings. However, apart from the specific exceptions mentioned above, it cannot be argued that there are overall leader and laggard sectors and the most represented sectors, such as the Industrials, stretch across the whole classification. A distinctive feature of SMEs is that mostly originate from manufacturing industries; although there is a broad volume of studies discussing the sustainability performance of SMEs (Malesios et al., 2021), we find limited quantitative assessments at sectoral level in

the literature.

Altogether, due to both the imputation procedure and the outranking nature of the model, the degree of voluntary disclosure achieved by European companies reporting on sustainability issues is definitely a crucial determinant of rankings.

Recall that we implicitly assume that it is important for a firm to be consistent across the three sustainability dimensions, irrespective of sectoral peculiarities, by attributing equal importance to each sustainability dimension of the firm. This explains in part why some companies that may be focused on reporting on a specific sustainability dimension, which is deemed to be more material than others, tend to underperform.

Given the large number of real estate firms reporting ESG data and the significant heterogeneity in rankings, we briefly focus on such firms to better understand the drivers of their ESG performance. In addition, some important distinguishing traits of real estate companies with regards to ESG performance have been recently documented, involving, in particular, the (positive) relationship between sustainability achievements and occupancy rates, property prices, and debt financing (Feng and Wu, 2023). With respect to our results, in Table 7 a breakdown of the scores is proposed. High variability and lack of consistency emerge across the three sustainability dimensions. Furthermore, poor or insufficient track records with respect to specific pillars emerge for a few companies, which receive scores close to zero. Therefore, also in this case, we observe much lower scores for both the social and the governance pillars, whereas the worst-ranked firm does not actually report on environmental issues.

Our results are definitely grounded in the specific and crucial assumption introduced in this section, according to which all the ESG profiles are equally material: it follows that more emphasis is put on the ability of a company to perform well across all the dimensions.

Nonetheless, as we discuss in Section 5.2, the performance of a few companies is found to be relatively insensitive to different parameterizations. Also, for the remaining companies, by letting the materiality of ESG profiles vary randomly, we do not observe extreme variations in the final ordering of firms. Similar conclusions are reached by Barro et al. (2024), where a sector-neutral analysis is conducted in a stochastic multiacceptability analysis (SMAA) framework, showing that the best-performing SMEs (on average), tend to remain the preferred choice even when taking into account very different preference settings. Such a pattern is worthy of remark since it allows policymakers to capture exactly which firms or sectors lead or lag behind on sustainable transition in a consistent way, and to design targeted interventions accordingly. Furthermore, stable rankings are valuable also for

Table 7: A focus on real estate firms. The overall score is reported along with a breakdown by pillar. Note that, by construction, the overall score based on the net flow φ_{net} is non-compensatory.

Ranking	ID	Sector	Country	ESG	E	S	G
16	51	Real Estate	Belgium	59.68	73.17	76.38	28.45
31	88	Real Estate	Sweden	52.50	63.20	44.80	61.11
42	76	Real Estate	Poland	49.26	58.31	64.76	27.69
46	95	Real Estate	Sweden	48.11	63.20	3.89	54.07
59	35	Real Estate	Denmark	43.80	52.84	1.00	50.33
64	101	Real Estate	Sweden	41.26	56.37	2.61	47.46
65	96	Real Estate	Swedeny	41.15	56.37	13.17	46.90
67	47	Real Estate	Estonia	40.70	29.18	68.17	52.69
69	71	Real Estate	France	40.06	43.61	69.65	34.52
79	99	Real Estate	Sweden	34.71	49.31	32.61	36.06
86	78	Real Estate	Poland	30.13	36.34	39.83	3.18
88	100	Real Estate	Sweden	29.16	39.47	13.12	31.71
97	13	Real Estate	Italy	20.88	0.75	16.56	37.93

firms, as this provides an indication of little disagreements e.g. across different aggregation and elicitation choices by rating agencies, which is a well-known problem both in the ESG literature and in practice (Billio et al., 2021).

5.2 Sensitivity analysis

In what follows, we aim to document the impact of variations across the parameter space, to verify how the underlying model assumptions affect both the robustness of the preference ordering and the scores. In what follows, we analyze the sensitivity of the model with reference to two sources of uncertainty: different parameter settings and the imputation procedure.

5.2.1 Assessing the robustness of the model parameterization

We employ an All-(factors)-At-a-Time (AAT) approach (Pianosi et al., 2016), in order to assess the model sensitivity to parameter variations. In AAT methods, output variations are induced by altering all the input factors simultaneously. We assume that all the input factors are simultaneously drawn from independent uniform random variables.

We perform $s = 10,000$ Monte Carlo simulations to identify promising regions of the inputs space. After testing the MURAME across different ranges of thresholds and for all the admissible values in the feasible region of weights, we report the results for reasonably

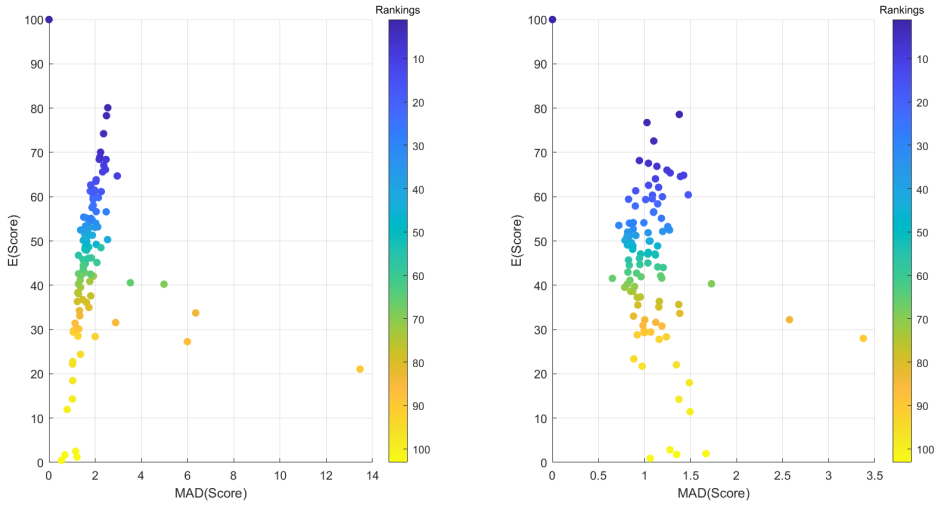

Figure 2: Sensitivity of firms' scores to thresholds settings (left panel) and across different weighting schemes (right panel) in a Mean-MAD framework. Colors in both panels are used to denote the rankings attributed to different alternatives.

broad ranges of values, as follows:

- $w_i \in [0, 1]$, s.t. $\sum_i^n w_i = 1$ and $w_i \geq 0 \quad \forall i$;
- The indifference threshold q_i is generated in a range between the 15th and the 25th percentile, the preference thresholds p_i between the 55th and the 65th percentile and finally the veto threshold v_i between 75th and the 85th percentile of criterion i.

Some preliminary tests show that the impact of weight settings on scores and rankings is mostly firm-specific and somewhat sector-specific. No country-specific clustering is observed. As for the stability of the scores, we find that top performers are quite robust to perturbations in weights, whereas we observe an increase in mean-absolute deviation (MAD) for poorly performing firms; overall, MAD is found to be low. With respect to the threshold settings, specific clustering effects across countries and sectors are less clear. In this case, top performers seem to be slightly more sensitive to variations in the value of the thresholds, although we observe a few outliers among worst-performing firms with a large mean-absolute deviation. The results are reported in Figures 2, 3, 4.

We also document the properties of rating assignments by assessing two basic measures of uncertainty and robustness of the assignments. Following (Doumpos and Figueira, 2019),

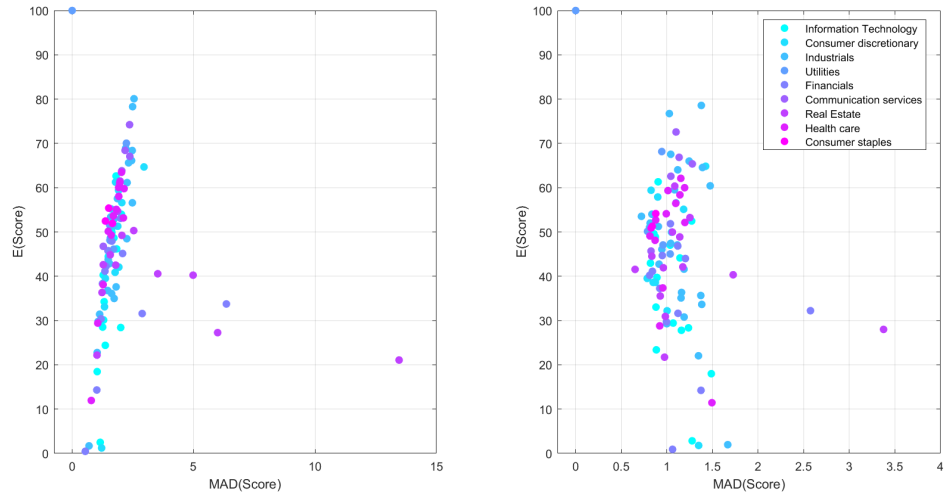

Figure 3: Sectoral clusters of firms with respect to sensitivity to thresholds settings (left panel) and weights settings (right panel) are reported, in a Mean-MAD framework. Roughly, some firms have a clustering tendency. Note that the number of observations across different sectors may vary significantly.

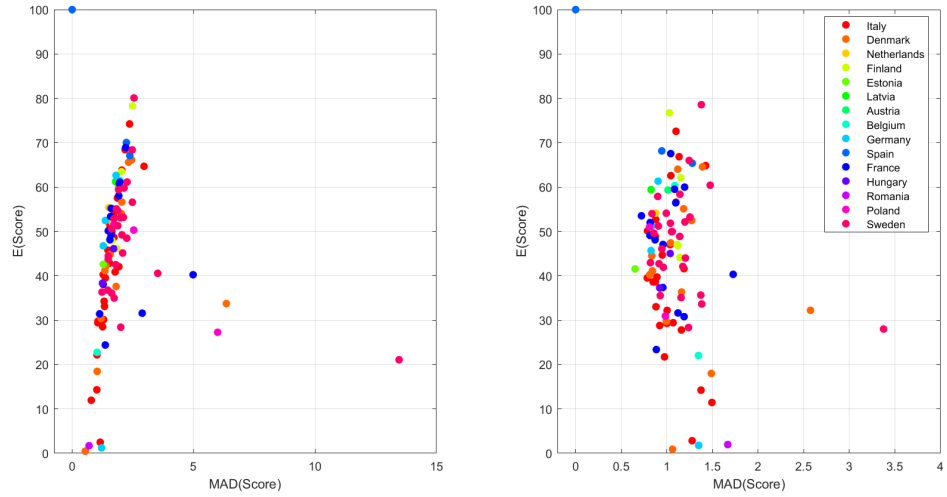

Figure 4: Geographical clusters of firms with respect to sensitivity to thresholds (left panel) and weights (right panel) are reported, in a Mean-MAD framework. No country effects emerge from data. Note that the number of observations across different countries may vary significantly.

Table 8: Average of the range and entropy assignments across various feasible values of $w \in [0,1]$ for different percentile-based specifications of the thresholds. The selected threshold settings used in the application are reported in bold.

Thresholds			Uncertainty Measures	
q	p	v	Entropy	Range
0	0.5	0.7	0.34	11.41
0	0.5	1	0.55	29.59
0.2	**0.6**	**0.8**	**0.34**	**11.08**
0.25	0.75	1.5	0.81	80.99
0.5	1	2	0.79	78.16

we compute first the range of the assignments $\bar{\mathcal{R}} \in [0, 103]$ w.r.t. the attained rankings $k = 1, \ldots, m$, with $m = 104$, across a set of simulations of weights, for a given specification of thresholds, of size $s = 10,000$. It corresponds to the average range across all m firms, with $\mathcal{U}$ and $\mathcal{L}$ denoting respectively the maximum and the minimum ranking for a given company i over s simulations:

$$\bar{\mathcal{R}} = \frac{1}{n} \sum_{i=1}^{m} (\hat{\mathcal{U}}_i - \hat{\mathcal{L}}_i) \tag{7}$$

We also compute the entropy $\bar{\mathcal{E}} \in [0,1]$ of the assignments w.r.t. the attained rankings $k = 1, \ldots, m$ across s simulations, where p_{ik}^s is the percentage of simulations w.r.t. different assignments: a low entropy indicates a robust ordering across all simulations s of weights, for a given specification of thresholds; high entropy indicates instead a high level of variability in the results:

$$\bar{\mathcal{E}} = \frac{1}{m} \sum_{i=1}^{m} \Big[- \frac{1}{ln(m)} \sum_{k=1}^{m} p_{ik}^s ln(p_{ik}^s) \Big] \tag{8}$$

Table 8 summarize the results for five different settings of the thresholds, including some extreme combinations of indifference, preference and veto thresholds and we document that, as long as settings are specified within reasonable ranges, a low degree of uncertainty and high robustness can be attained. In particular, the third specification seems to us a natural choice for being relatively neutral.

To assess the influence of both weights and thresholds, we also plot a stacked area chart in Figure 5, where each color denotes the probability of a specific firm ending up in different positions of the overall rankings. The results are derived by simulating randomly

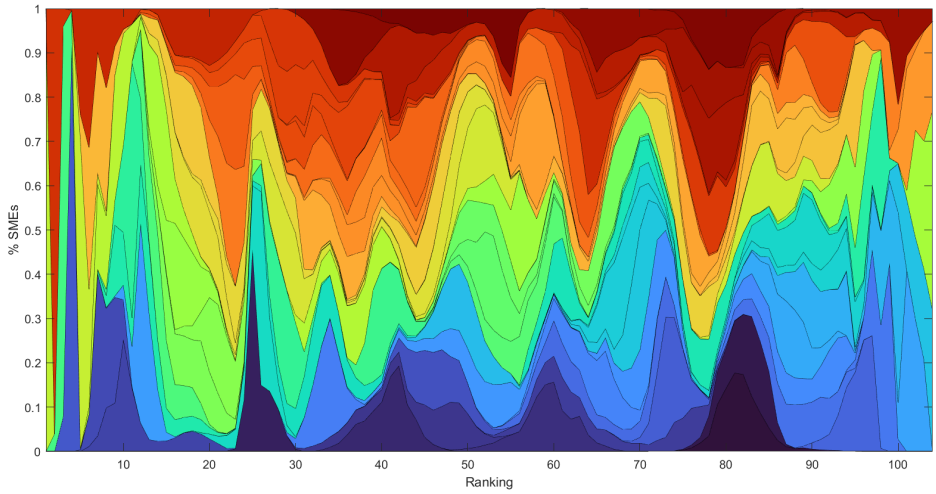

Figure 5: Stacked area chart of firms' weights in the general standings across different settings. Note that on the x-axis the ranking for each company across $s = 10,000$ simulations is reported, while on the y-axis the probability of a SMEs ending up in a given ranking is represented with a color for each SME.

the variation of all parameters jointly. High concentration of areas vertically denotes a stable ranking for that specific firm. Note that such behaviour tends to emerge both on the left and the right of the chart, whereas intermediate rankings are more volatile and erratic. Intuitively, this implies that both top and bottom companies are also more likely to respectively outperform and underperform competitors from a sustainability point of view companies under different scenarios.

5.2.2 Assessing the robustness of the imputation procedure

In what follows, we gauge the impact of the proposed imputation procedure. Our point for making a further robustness check is that a replacement of missing values with a by-sector point estimate based on the worst value, might make pairwise comparisons between alternatives insensitive across different threshold settings. In this way, it could be supposed that the proposed approach might artificially induce robustness in the rankings and scores across different parameter settings.

Therefore, to assess the effectiveness of our imputation choice, we generate replacements for missing values according to a multiple imputation procedure. We impute missing values

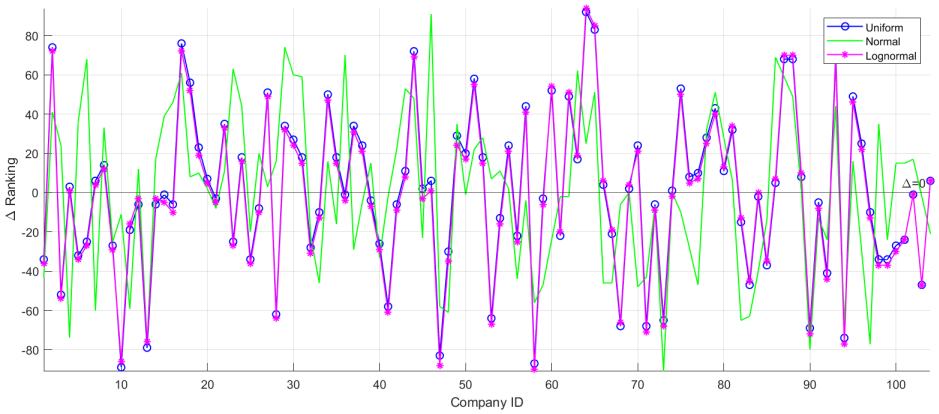

Figure 6: Each coloured plot denotes the differences between the reference ranking (see the horizontal line in black with $\Delta = 0$) and the rankings derived with multiple imputation methods.

by sampling respectively from a uniform, a normal, and a lognormal distribution 10,000 times. We assume prudentially that missing data are sampled from a distribution fitted on below (above)-median observed data, depending respectively whether the criteria is to be maximized or minimized.

A two-sample Kolmogorov-Smirnov test is applied to check whether each distribution of scores, whose missing values have been reconstructed with a multiple imputation method, is statistically different from the reference distribution of scores based on imputation with the worst sector value. The null hypothesis cannot be rejected for all the three tests.

Finally, in Figure 6 the difference between the reference ranking ($\Delta = 0$) and the rankings based on multiple imputation is reported, with notable differences for most companies, apart from a few exceptions for firms placed on the right tail of the distribution (see in particular IDs 8, 9, 11, 42, 56, 57, 70 and 91).

We conclude that, although the variability of the rankings is not negligible, top firms are not only robust to different settings, but also to various imputation methods.

5.3 Assessing the impact of a prudential single imputation procedure

In accordance with the approximation that lack of data might imply unwillingness or inability to release information, we impute missing data in a given sector with the sectoral worst value. Such a heuristic replacement rule entails a variety of statistical and economic

consequences, as discussed in Section 4.2. However, here we further elaborate on this point, as we factor in other model-specific consequences. First, the computation of inflows and outflows in an outranking approach is specifically affected by lack of disclosure, leading to system-wide effects as the *true* values of missing observations might actually be better or worse than estimated and ultimately impacting on the overall scoring methodology. In Proposition 5.1 we formally show that, given a decision matrix $\mathbf{G}_{m\times n}$, where the unknown but true quantities of interest are denoted with $\tilde{(\cdot)}$ and with exactly one missing value w.r.t. a criterion-alternative pair, it is always convenient for firm i to disclose the (not publicly known) true observation in place of the missing one, in the sense that the alternative is at least as good as if the value were unknown, so that ESG disclosure turns out to be the optimal strategy that a firm should pursue.

Proposition 5.1 *Let a_i be the only alternative among the m ones for which its performance in relation to the j-th criterion is unknown (so, g_{ij} is prudentially set to 0), let $\tilde{g}_{ij} \geq 0$ be the unknown true performance of a_i before transformation (6), and let $\tilde{\varphi}(a_i)$ be the net flow of a_i computed considering the unknown but true performance of a_i w.r.t. the j-th criterion. If there exist at least an alternative a_k, with $k \neq i$, such that*

- *$g_{kj} > 0$ in case the j-th criterion is to be maximized,*
- *$g_{kj} < 1$ in case the j-th criterion is to be minimized*

then $\varphi(a_i) \leq \tilde{\varphi}(a_i)$.

Proof. See Appendix A ■

Moreover, in Proposition 5.2 we extend Proposition 5.1 to the case of two or more missing observations for the same criterion, for two or more alternatives, and we discuss why there are no optimal strategies that can be pursued by firms, since the "true" net flow depends also on pairwise comparisons between unknown true values. Our findings show that, in case of two or more companies failing to report on sustainability topics, undesired outcomes might occur.

Proposition 5.2 *Let $a_{i_1}, a_{i_2}, \ldots, a_{i_p}$, with $\#\{i_1, i_2, \ldots, i_p\} \in \{2, 3, \ldots, m-1\}$ be the alternatives, among the m ones, whose performance in relation to the j-th criterion are unknown (so, $g_{i_1 j}, g_{i_2 j}, \ldots, g_{i_p j}$ are prudentially set to 0), let $\tilde{g}_{i_1 j} \geq 0, \tilde{g}_{i_2 j} \geq 0, \ldots, \tilde{g}_{i_p j} \geq 0$*

be the unknown performance of $a_{i_1}, a_{i_2}, \ldots, a_{i_p}$ *respectively before transformation* (6)*, and let* $\tilde{\varphi}(a_{i_1}), \tilde{\varphi}(a_{i_1}), \ldots, \tilde{\varphi}(a_{i_p})$ *be the net flows of* $a_{i_1}, a_{i_2}, \ldots, a_{i_p}$ *respectively computed considering the true but unknown performance of* $a_{i_1}, a_{i_2}, \ldots, a_{i_p}$ *w.r.t. the* j*-th criterion. If there exist at least an alternative* a_k *with* $k \notin \{i_1, \ldots, i_2, \ldots, i_p\}$ *such that*

- $g_{kj} > 0$ *in case the* j*-th criterion is to be maximized,*
- $g_{kj} < 1$ *in case the* j*-th criterion is to be minimized,*

then $\varphi(a_i) \gtreqless \tilde{\varphi}(a_i)$.

Sketch of proof. See Appendix A ■

Given that outranking approaches constitute a benchmark for complex scoring problems, important implications for policymakers can be derived from the results. Since no conclusions on optimality of disclosure strategies can be drawn, firms found to have poor ESG performance could possibly take advantage from underreporting their ESG performance, ultimately hindering the sustainable development goals of public policies. Therefore, we recommend that policymakers address such concerns, by discouraging opportunistic disclosure strategies and by appropriately establishing and monitoring minimum sustainability reporting standards.

6 Discussion

Based on our research findings, we observe that our scoring procedure has interesting connections with corporate theory. In this study, we make a proposal for capturing the the ESG performance of SMEs with a unique score; we conclude that the scoring system should account appropriately for data quality by aligning our figures to GRI standards and by setting disclosure requirements. We find that a subsample of firms actually perform robustly well across all dimensions. Our framework addresses policymakers and investors willing to grasp the current sustainability performance of SMEs, based on voluntary efforts, under limited assumptions about the reporting behaviour of firms. Our results are therefore valuable for future assessments of the ESG performance of SMEs. Currently, they compare with a mixed assessment on the role of ESG disclosure in the literature. According to (Gjergji et al., 2021), it has been found to have direct costs, but widely uncertain benefits. Other works stress that SMEs are better at adopting ESG practices compared to large firms (Ortiz-Martínez and Marín-Hernández, 2022), thanks to leaner structures, and that such

practices improve the competitiveness of firms (Torugsa et al., 2012). Propositions 5.1 and 5.2 generalize the discussion around the disclosure decision, by hypothesizing two settings where a firm is faced with the decision to release information. Under the most realistic setting with many firms not reporting data, we capture the uncertainty over the advantages of the disclosure decision, by allowing for a range of possible actions. Moreover, in Section 3 we find that the quality and the degree of disclosure is a country and sector-specific factor, which is a further relevant aspect previously investigated in the literature (Baldini et al., 2018), where it is stressed that social structures can play an important role in the definition of voluntary disclosure initiatives. Firms can also leverage ESG data to gain legitimacy by enhancing their corporate image (Gerwanski, 2020). Nonetheless, voluntary disclosure might still be specifically tailored to certain stakeholders' tastes, resulting in inconsistent reporting (O'Dochartaigh, 2019), as we highlight in Section 4.2. Our proposal also serves as a decision making tool for SMEs to channel information to relevant parties, as it can provide valuable information for several stakeholders, including policymakers and investors, and for the firm itself. First, it allows to derive its positioning, w.r.t. competitors and the supply chain. Sustainability disclosure can be therefore leveraged to improve corporate performance and enhance brand value and reputation, ultimately gaining a competitive advantage. Among other possible benefits, we mention also greater access to credit and better funding opportunities. In this respect, evidence in this direction is still limited to small samples (Nigri and Del Baldo, 2018; Lopez-Torres, 2023). Second, also in light of future mandatory requirements, within the framework presented in Section 5.3, we address some of the criticalities related to inconsistent reporting and we stress the importance for the policymaker to establish a level playing field, where clear and minimal standards are set. Unfortunately, current voluntary individual efforts are limited, possibly shaped by the demand of larger firms in the supply chain, in terms of sustainability reporting Rodríguez-Gutiérrez et al. (2021). The framework designed by the CSRD is therefore a necessary step to ensure that firms operate on equal footing; evidence in the literature shows indeed that country-level factors such as political systems, labor protection and social cohesion are key determinants shaping ESG disclosure practices and leading to reporting discrepancies (Baldini et al., 2018). The importance of a quantitative ESG assessment for the policymaker is also crucial for several other purposes, as it allows to monitor the current state of the ESG transition process, and provides insights into voluntary disclosure choices across countries and sectors, highlighting existing gaps in reporting standards; it also shows which firms are the frontrunners, allowing a better implementation of policies (i) targeting

a more transparent and balanced disclosure across different sectors, and (ii) providing the right incentives to firms, aiming at encouraging disclosure of more ESG data. Third, the ESG assessment procedure we devise has value for investors, as it increases transparency and reduces information asymmetries between the firm and markets, possibly supporting investment decisions and allowing a better understanding of all the risk profiles. The flipside is that monetizing the ESG implementation is much more difficult for SMEs. Therefore, Garrido-Ruso et al. (2024) suggest to establish simplified standards, a point we addressed throughout the paper, by proposing a flexible model, where complexity is kept at minimum. We believe that our results are also valuable also for the practitioner, since they contain insights for creditworthiness assessment of loan applicants. Although we do not discuss the implications for credit risk of our model, the next step is to understand whether the ESG dimension provides added value as a potential credit risk mitigation factor (Brogi et al., 2022) or whether the SMEs' carbon footprint, i.e. the exposure to climate risk, can predict a firm's default risk (Capasso et al., 2020), for which the literature has established a clear connection, also for the banking sector (Palmieri et al., 2024). In light of the quality of available data at the moment, it directly follows that little can be said w.r.t. SMEs at the moment; nonetheless, a key aspect requiring further investigation involves the inclusion of ESG metrics into credit risk ratings is equally important for predicting SMEs creditworthiness; otherwise, one should conclude that, in stark contrast with large companies, ESG disclosure might increase default risk, or be totally ineffective.

7 Conclusions

In this contribution we have analyzed the sustainability profiles in SMEs with a MCDA approach. We have set up a model where a limited number of assumptions is necessary, based on a robust percentile-based preference structure of the DM.

Firstly, we stressed the importance of identifying a set of variables that are relevant to firms and investors. To do this, we constructed a unique hand-collected dataset based on Global Reporting Initiative (2022)-compliant criteria, in order to make the assessment of results aligned with global standards for sustainability impact, and tailored to the distinctive features of SMEs. The obtained rankings were aimed at capturing leader and laggard firms in terms of ESGness after controlling for sector effects. Moreover, by adopting a prudential imputation approach, we dealt with the critical issue of missing ESG data for SMEs. A sensitivity analysis finally confirmed that our results are robust across different

model parameterizations and imputation methods. Since the replacement rule of missing data we use entails various economic consequences, we also proposed to assess how this might affect the firm's decision to actually release ESG data. Leveraging voluntary reporting, which has so far discouraged in-depth analysis of ESG practices in SMEs, our assessment constitutes a first step towards a comprehensive and systematic assessment of the ESG performance of SMEs, by addressing the well-known problem of transparency and comparability of the ESG scores (Billio et al., 2021), and providing a foundation for further analysis, with an eye to the interpretability of scores. The quality of data represents a major barrier in research on small business finance. Nonetheless, further research should be extended towards three directions. First, more granular databases would allow to break down data by industry and by country, allowing to characterize in detail the firms' sustainability performance and to tailor a suitable multicriteria approach accordingly (Doumpos et al., 2017). Moreover, by leveraging the time dimension, it would be possible to capture time-varying effects jointly with country and sector characteristics and to shed light more accurately on the ESG determinants and disclosure policies. A further improvement of our proposal can be attained by considering additional key performance indicators to develop a more comprehensive measure balancing conflicting criteria. Second, the integration of the ESG dimensions in credit risk models for SMEs would unlock a deeper understanding of its influence on firms' creditworthiness, starting from evidence in the literature pointing to the relevance of the ESG dimensions for large cap firms (Brogi et al., 2022). The importance of other non-financial information for SMEs has already been highlighted in a previous investigation (Altman et al., 2010), possibly suggesting that valuable information can be extracted from ESG data for credit analysts as well. Third, the materiality assessment is a crucial point briefly mentioned in this work requiring further elaboration. Previous findings have established a connection between the degree of stakeholder engagement and the materiality analysis process (Manetti, 2011; Torelli et al., 2020), but do not provide a modeling framework. Future work, following e.g. (Corazza et al., 2015), should provide a framework for deriving the materiality of issues, based on a preference disaggregation approach, so that the stakeholders' preferences can be inferred from actual evaluations of alternatives.

Disclaimer

This paper was developed within the projects funded by: Next Generation EU - "GRINS - Growing Resilient, INclusive and Sustainable" project (PE0000018), National Recovery and Resilience Plan (NRRP) – PE9 - Mission 4, C2, Intervention 1.3; Next Generation EU - "Just Energy Transition – JET: Stochastic and machine learning methods for the evaluation, mitigation and geographical hedging of involved natural risks (with climate in view)" project (P2022XTLM2), National Recovery and Resilience Plan (NRRP) - Mission 4, C2, Intervention 1.1. The views and opinions expressed are only those of the authors and do not necessarily reflect those of the European Union or the European Commission or the Italian Ministry of University and Research. Neither the European Union, the European Commission nor the Italian Ministry of University and Research can be held responsible for them.

Acknowledgments

We would like to thank participants at the Workshop of Sustainable Finance, University of Venice, and of the 96th meeting of the EURO Working Group on Multiple Criteria Decision Aiding.

A Appendix: Proofs of Propositions

Proof of Proposition 5.1. Let us focus on the case where the j-th criterion is to be maximized. The fact that $\tilde{g}_{ij} \geq 0$ implies the following regarding the local concordances and local discordances, for the pairs of alternatives (a_i, a_k) and (a_k, a_i) w.r.t. the j-th criterion:

- $C_j(a_i, a_k) \leq 1$ by definition, and $C_j(a_k, a_i) = 1$ since $g_{ij} - g_{kj} = -g_{kj} < 0 < q_j$ (see Equation (1));
- $\tilde{C}_j(a_i, a_k) \geq C_j(a_i, a_k)$ as $g_{kj} - \tilde{g}_{ij} \leq g_{kj} - g_{ij} = g_{kj}$ and $\tilde{C}_j(a_k, a_i) \leq C_j(a_k, a_i)$ since $\tilde{g}_{ij} - g_{kj} \geq g_{ij} - g_{kj} = -g_{kj}$ (see Equation (1));
- $D_j(a_i, a_k) \geq 0$ by definition, and $D_j(a_k, a_i) = 0$ since $g_{ij} - g_{kj} = -g_{kj} < 0 < p_j$ (see Equation (2));

- $\tilde{D}_j(a_i, a_k) \leq D_j(a_i, a_k)$ as $g_{kj} - \tilde{g}_{ij} \leq g_{kj} - g_{ij} = g_{kj} = g_{kj}$, and $\tilde{D}_j(a_k, a_i) \geq D_j(a_k, a_i)$ since $\tilde{g}_{ij} - g_{kj} \geq g_{ij} - g_{kj} = -g_{kj}$ (see Equation (2)).

So, concerning the global concordances of pairs (a_i, a_k) and (a_k, a_i), one has:

$$\tilde{C}(a_i, a_k) = \sum_{\substack{l=1 \\ l \neq j}}^{n} w_l C_l(a_i, a_k) + w_j \tilde{C}_j(a_i, a_k) \geq C(a_i, a_k) = \sum_{\substack{l=1 \\ l \neq j}}^{n} w_l C_l(a_i, a_k) + w_j C_j(a_i, a_k) \tag{9}$$

since $\tilde{C}_j(a_i, a_k) \geq C_j(a_i, a_k)$ and

$$\tilde{C}(a_k, a_i) = \sum_{\substack{l=1 \\ l \neq j}}^{n} w_l C_l(a_k, a_i) + w_j \tilde{C}_j(a_k, a_i) \leq C(a_k, a_i) = \sum_{\substack{l=1 \\ l \neq j}}^{n} w_l C_l(a_k, a_i) + w_j C_j(a_k, a_i) \tag{10}$$

as $\tilde{C}_j(a_k, a_i) \leq C_j(a_k, a_i)$.

Similarly, concerning the outranking indices of the pairs (a_i, a_k) and (a_k, a_i) one has

$$\tilde{O}(a_i, a_k) \geq O(a_i, a_k) \tag{11}$$

since $\tilde{C}_j(a_i, a_k) \geq C_j(a_i, a_k)$ and $\tilde{D}_j(a_i, a_k) \leq D_j(a_i, a_k)$, and one has

$$\tilde{O}(a_k, a_i) \leq O(a_k, a_i) \tag{12}$$

as $\tilde{C}_j(a_k, a_i) \leq C_j(a_k, a_i)$ and $\tilde{D}_j(a_k, a_i) \geq D_j(a_k, a_i)$.

Lastly, with reference to the net flows $\varphi(a_i) = \varphi^+(a_i) - \varphi^-(a_i)$ and $\tilde{\varphi}(a_i) = \tilde{\varphi}^+(a_i) - \tilde{\varphi}^-(a_i)$ (see Equation (5)), from the above one has that all the addends of $\varphi^+(a_i)$ are lower or equal than the corresponding addends of $\tilde{\varphi}^+(a_i)$ and that all the addends of $\varphi^-(a_i)$ are greater or equal than the corresponding addends of $\tilde{\varphi}^-(a_i)$, so:

$$\varphi^+(a_i) \leq \tilde{\varphi}^+(a_i) \quad \text{and} \quad \tilde{\varphi}^-(a_i) \geq \tilde{\varphi}^-(a_i) \tag{13}$$

therefore

$$\varphi(a_i) = \varphi^+(a_i) - \varphi^-(a_i) \leq \tilde{\varphi}(a_i) = \tilde{\varphi}^+(a_i) - \tilde{\varphi}^-(a_i) \tag{14}$$

This proves the thesis. Similarly, one can prove the case where the j-th criterion is to be minimized. ■

Sketch of the proof of Proposition 5.2. Let us focus again on the case where the j-th criterion is to be maximized. Concerning the comparison of pairs of alternatives in which one alternative has an unknown performance and the other alternative has a known performance (as in Proposition 5.1), all the inequalities proved in Proposition 5.1 hold again. But as for the comparison of pairs of alternatives in which both the alternatives have unknown performance, none of the inequalities proved in Proposition 5.1 hold in general anymore. Therefore, the contributions of the net flows of the latter kind of comparisons cannot be further evaluated and, consequently, it is no longer possible to define an ordering between $\phi(a_i)$ and $\tilde{\phi}(a_i)$. Similarly, one can prove the case where the j-th criterion is to be minimized. ■

References

Altman, E. I. and Sabato, G. (2007). Modelling credit risk for SMEs: Evidence from the US market. *Abacus*, 43(3):332–357.

Altman, E. I., Sabato, G., and Wilson, N. (2010). The value of non-financial information in small and medium-sized enterprise risk management. *The Journal of Credit Risk*, 6(2):95.

Angilella, S. and Mazzù, S. (2015). The financing of innovative SMEs: A multicriteria credit rating model. *European Journal of Operational Research*, 244(2):540–554.

Baier, P., Berninger, M., and Kiesel, F. (2020). Environmental, social and governance reporting in annual reports: A textual analysis. *Financial Markets, Institutions & Instruments*, 29(3):93–118.

Baldini, M., Maso, L. D., Liberatore, G., Mazzi, F., and Terzani, S. (2018). Role of country- and firm-level determinants in environmental, social, and governance disclosure. *Journal of Business Ethics*, 150:79–98.

Barro, D., Corazza, M., and Filograsso, G. (2024). A robust sustainability assessment for smes based on multicriteria decision aiding. In Corazza, M., Gannon, F., Legros, F., Pizzi, C., and Touzé, V., editors, *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, pages 43–48. Springer Nature Switzerland.

Bartolacci, F., Caputo, A., and Soverchia, M. (2020). Sustainability and financial performance of small and medium sized enterprises: A bibliometric and systematic literature review. *Business Strategy and the Environment*, 29(3):1297–1309.

Baumüller, J. and Grbenic, S. O. (2021). Moving from non-financial to sustainability reporting: Analyzing the EU commission's proposal for a Corporate Sustainability Reporting Directive (CSRD). *Facta Universitatis, Series: Economics and Organization*, (1):369–381.

Billio, M., Costola, M., Hristova, I., Latino, C., and Pelizzon, L. (2021). Inside the ESG

ratings:(Dis) agreement and performance. *Corporate Social Responsibility and Environmental Management*, 28(5):1426–1445.

Bofinger, Y., Heyden, K. J., and Rock, B. (2022). Corporate social responsibility and market efficiency: Evidence from esg and misvaluation measures. *Journal of Banking & Finance*, 134:106322.

Brans, J.-P. and Vincke, P. (1985). Note—A Preference Ranking Organisation Method: (The PROMETHEE Method for Multiple Criteria Decision-Making). *Management Science*, 31(6):647–656.

Brogi, M., Lagasio, V., and Porretta, P. (2022). Be good to be wise: Environmental, Social, and Governance awareness as a potential credit risk mitigation factor. *Journal of International Financial Management & Accounting*, 33(3):522–547.

Capasso, G., Gianfrate, G., and Spinelli, M. (2020). Climate change and credit risk. *Journal of Cleaner Production*, 266:121634.

Corazza, M., Funari, S., and Gusso, R. (2015). An evolutionary approach to preference disaggregation in a MURAME-based creditworthiness problem. *Applied Soft Computing*, 29:110–121.

Corazza, M., Funari, S., and Gusso, R. (2016). Creditworthiness evaluation of Italian SMEs at the beginning of the 2007–2008 crisis: An MCDA approach. *The North American Journal of Economics and Finance*, 38:1–26.

Czarnitzki, D. and Hottenrott, H. (2011). R&D investment and financing constraints of small and medium-sized firms. *Small Business Economics*, 36:65–83.

De Falco, S. E., Montera, R., Leo, S., Laviola, F., Vito, P., Sardanelli, D., Basile, G., Nevi, G., and Alaia, R. (2024). Trends and patterns in esg research: A bibliometric odyssey and research agenda. *Corporate Social Responsibility and Environmental Management*.

Deshpande, P. C., Skaar, C., Brattebø, H., and Fet, A. M. (2020). Multi-criteria decision analysis (MCDA) method for assessing the sustainability of end-of-life alternatives for waste plastics: A case study of Norway. *Science of the total environment*, 719:137353.

Doumpos, M., Andriosopoulos, K., Galariotis, E., Makridou, G., and Zopounidis, C. (2017). Corporate failure prediction in the European energy sector: A multicriteria approach and the effect of country characteristics. *European Journal of Operational Research*, 262(1):347–360.

Doumpos, M. and Figueira, J. R. (2019). A multicriteria outranking approach for modeling corporate credit ratings: An application of the Electre Tri-nC method. *Omega*, 82:166–180.

D'Amato, A. (2020). Capital structure, debt maturity, and financial crisis: empirical evidence from SMEs. *Small Business Economics*, 55(4):919–941.

El Ghoul, S., Guedhami, O., Kwok, C. C., and Mishra, D. R. (2011). Does corporate social responsibility affect the cost of capital? *Journal of Banking & Finance*, 35(9):2388–2406.

ESRS (2023). European sustainability reporting standards. Available at `https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/`. Accessed: 2023-06-30.

European Commission (2023). Corporate sustainability reporting. Available at `https://eur-lex.europa.eu/legal-content/`. Accessed: 2023-06-30.

Feng, Z. and Wu, Z. (2023). ESG disclosure, REIT debt financing and firm value. *The Journal of Real Estate Finance and Economics*, 67(3):388–422.

Figueira, J. R., Mousseau, V., and Roy, B. (2016). ELECTRE methods. In *Multiple criteria decision analysis: State of the art surveys*, International Series in Operations Research & Management Science, S. Greco, M. Ehrgott and J. Figuiera, pages 155–185. Springer.

Freeman, R. E. and Reed, D. L. (1983). Stockholders and stakeholders: A new perspective on corporate governance. *California management review*, 25(3):88–106.

Friede, G., Busch, T., and Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4):210–233.

Garrido-Ruso, M., Otero-González, L., López-Penabad, M.-C., and Santomil, P. D. (2024). Does ESG implementation influence performance and risk in SMEs? *Corporate Social Responsibility and Environmental Management*.

Gerwanski, J. (2020). Managers' incentives and disincentives to engage with integrated reporting or why managers might not adopt integrated reporting: an exploratory study in a nascent setting. *Qualitative Research in Accounting & Management*, 17(4):553–587.

Giese, G., Lee, L.-E., Melas, D., Nagy, Z., and Nishikawa, L. (2019). Foundations of ESG investing: How ESG affects equity valuation, risk, and performance. *The Journal of Portfolio Management*, 45(5):69–83.

Gjergji, R., Vena, L., Sciascia, S., and Cortesi, A. (2021). The effects of environmental, social and governance disclosure on the cost of capital in small and medium enterprises: The role of family business status. *Business Strategy and the Environment*, 30(1):683–693.

Global Reporting Initiative (2022). Sustainability reporting guidelines. Available at `https://www.globalreporting.org/how-to-use-the-gri-standards/gri-standards-english-language/`. Accessed: 2023-06-30.

Goletsis, Y., Psarras, J., and Samouilidis, J.-E. (2003). Project ranking in the Armenian energy sector using a multicriteria method for groups. *Annals of Operations Research*, 120:135–157.

Goyal, P., Rahman, Z., and Kazmi, A. A. (2013). Corporate sustainability performance and firm performance research: Literature review and future research agenda. *Management Decision*, 51(2):361–379.

Guney, Y., Hernandez-Perdomo, E., and Rocco, C. M. (2020). Does relative strength in corporate governance improve corporate performance? Empirical evidence using MCDA approach. *Journal of the Operational Research Society*, 71(10):1593–1618.

Henriksson, R., Livnat, J., Pfeifer, P., and Stumpp, M. (2019). Integrating ESG in portfolio construction. *The Journal of Portfolio Management*, 45(4):67–81.

Höck, A., Klein, C., Landau, A., and Zwergel, B. (2020). The effect of environmental sustainability on credit risk. *Journal of Asset Management*, 21(2):85–93.

Kotsantonis, S. and Serafeim, G. (2019). Four things no one will tell you about ESG data. *Journal of Applied Corporate Finance*, 31(2):50–58.

Lindsey, L. A., Pruitt, S., and Schiller, C. (2022). The cost of ESG Investing. *Available at SSRN 3975077*.

Lopez-Torres, G. C. (2023). The impact of SMEs' sustainability on competitiveness. *Measuring business excellence*, 27(1):107–120.

Malesios, C., De, D., Moursellas, A., Dey, P. K., and Evangelinos, K. (2021). Sustainability performance analysis of small and medium sized enterprises: Criteria, methods and framework. *Socio-Economic Planning Sciences*, 75:100993.

Manetti, G. (2011). The quality of stakeholder engagement in sustainability reporting: empirical evidence and critical points. *Corporate social responsibility and environmental management*, 18(2):110–122.

Moro, A. and Fink, M. (2013). Loan managers' trust and credit access for SMEs. *Journal of Banking & Finance*, 37(3):927–936.

Nigri, G. and Del Baldo, M. (2018). Sustainability reporting and performance measurement systems: How do small-and medium-sized benefit corporations manage integration? *Sustainability*, 10(12):4499.

Ortiz-Martínez, E. and Marín-Hernández, S. (2022). European SMEs and non-financial information on sustainability. *International Journal of Sustainable Development & World Ecology*, 29(2):112–124.

Ortiz-Martínez, E. and Marín-Hernández, S. (2023). Sustainability Information in European Small-and Medium-Sized Enterprises. *Journal of the Knowledge Economy*, pages 1–26.

O'Dochartaigh, A. (2019). No more fairytales: A quest for alternative narratives of sustainable business. *Accounting, Auditing & Accountability Journal*, 32(5):1384–1413.

Palmieri, E., Ferilli, G. B., Altunbas, Y., Stefanelli, V., and Geretto, E. F. (2024). Business model and ESG pillars: The impacts on banking default risk. *International Review of Financial Analysis*, 91:102978.

Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B., and Wagener, T. (2016). Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, 79:214–232.

Prencipe, A. (2004). Proprietary costs and determinants of voluntary segment disclosure: evidence from Italian listed companies. *European Accounting Review*, 13(2):319–340.

Rodríguez-Gutiérrez, P., Guerrero-Baena, M. D., Luque-Vílchez, M., and Castilla-Polo, F. (2021). An approach to using the best-worst method for supporting sustainability reporting decision-making in SMEs. *Journal of Environmental Planning and Management*, 64(14):2618–2640.

Sahin, Ö., Bax, K., Czado, C., and Paterlini, S. (2022). Environmental, Social, Governance scores and the Missing pillar—Why does missing information matter? *Corporate Social Responsibility and Environmental Management*, 29(5):1782–1798.

Shehata, N., Salhin, A., and El-Helaly, M. (2017). Board diversity and firm performance:

evidence from the UK SMEs. *Applied Economics*, 49(48):4817–4832.

Singhania, M. and Gupta, D. (2023). Impact of environmental, social and governance (esg) disclosure on firm risk: A meta-analytical review. *Corporate Social Responsibility and Environmental Management*.

Torelli, R., Balluchi, F., and Furlotti, K. (2020). The materiality assessment and stakeholder engagement: A content analysis of sustainability reports. *Corporate Social Responsibility and Environmental Management*, 27(2):470–484.

Torugsa, N. A., O'Donohue, W., and Hecker, R. (2012). Capabilities, proactive CSR and financial performance in SMEs: Empirical evidence from an Australian manufacturing industry sector. *Journal of Business Ethics*, 109:483–500.

Tsikriktsis, N. (2005). A review of techniques for treating missing data in OM survey research. *Journal of Operations Management*, 24(1):53–62.

Voulgaris, F., Doumpos, M., and Zopounidis, C. (2000). On the evaluation of Greek industrial SME's performance via multicriteria analysis of financial ratios. *Small Business Economics*, 15:127–136.

Zionts, S. (1979). MCDM—if not a roman numeral, then what? *Interfaces*, 9(4):94–101.