Green performance through the lens of the EU Taxonomy: Early evidence

Abstract

After the stepwise introduction of the EU Taxonomy Regulation in 2021, this paper is one of the first to analyse firm reported taxonomy data. We examine how well the new regulation captures green performance and compare taxonomy-based metrics to ESG ratings. We explore whether the taxonomy-based metrics can be used as an alternative to ESG ratings. We hand-collect taxonomy data for French, German, and Italian public firms in the years 2022-2023. We show that 36% of firms report no taxonomy-eligible or taxonomy-aligned activities. We further calculate green KPIs based on Alessi et al. (2024) and combine them into a novel Green Score. We investigate the associations between green KPIs (Green Score) and ESG ratings (E ratings) and find weak associations. We also document that the Green Score is positively related to firm size, profitability and capital intensity. Our findings suggest that ESG ratings (E ratings) fail to capture green performance and green activities as defined by the EU Taxonomy. The forward-looking and standardised taxonomy-based metrics seem to offer a promising path toward more transparent and comparable sustainability performance metrics.

1. Introduction

Sustainability has been at the forefront of international politics and has become part of the zeitgeist in the general public. While there is agreement that practices must change, progress in this area has been slow because of disagreements on decision-making levels, lobbying, and profit orientation. Therefore, regulators and firms are increasingly pressured to implement more transparent reporting practices and guide our planet toward a more sustainable future.

For this purpose, the European Commission (hereafter EC) agreed upon the European Green Deal in 2020, which builds on a set of actions that should help the EU become climate neutral by 2050 (EC, 2023). One of the EC's goals was to create comparable metrics across firms that measure the extent and effectiveness of sustainable investments for firms and investors. This led to the EU Taxonomy Regulation (EU Regulation 2020/852), that defines three green metrics that firms must report. The first objective of the EU Taxonomy is creating a frame of reference for investors and firms, and the second is supporting firms in their efforts to plan and finance the transition. This is achieved through clearly defined activities and corresponding requirements that make them sustainable. The EU Taxonomy also aims to protect against greenwashing practices by requiring comparable and transparent metrics and to accelerate financing sustainable projects and those needed in the transition (EC, 2024).

With sustainability becoming increasingly important for investors and firms in times of non-financial reporting, interested parties would often use ESG (environmental, social and governance) ratings as sustainability metrics. These ratings are provided by independent vendors and aggregate ESG activities of a firm in a few simple metrics. However, recent academic literature showed that ESG ratings are subjective and non-transparent, hence trust in them declined. We exploit the newly introduced EU Taxonomy and investigate the relationship between the ESG ratings and the taxonomy-based green metrics reported by firms from France, Germany and Italy. Following Alessi et al. (2024), we calculate three taxonomy- based key performance indicators (hereafter KPIs) and then combine them into our new green metric, the Green Score. We empirically test the association between the Green Scores and ESG ratings, with particular focus on the E pillar of the ratings (hereafter E rating). We propose the taxonomy-based Green Score as an alternative metric for assessing the sustainability of firms, portfolios and investments. We tackle the research question how effectively does the EU Taxonomy measure sustainability performance and how does it relate to ESG ratings.

We employ a sample of 91 French, German and Italian non-financial firms and hand collect their taxonomy data (i.e., eligible and aligned) for years 2022-2023. We document that only a subset of firms (64% of sample firms) reports metrics for taxonomy-eligible and taxonomy-aligned activities. With respect to green metrics, on average, 31% (9%) of revenue is reported as eligible (aligned), 46% (18%) of CapEx and 35% (15%) of OpEx is reported as eligible (aligned), respectively. We predict and find that the Green Score is weakly positively associated with E(SG) rating. The association is stronger for a subsample of high-quality reporters. Similarly, we find a positive association between Green Profit and E rating, which is stronger than the former. Moreover, our regression analysis shows that the Green Score is positively associated with firm size, profitability and capital intensity. It also shows an increase of the score for Italian and German firms compared to French ones. In sum, we interpret our findings as E(SG) ratings failing to capture green performance and green activities as defined by the EU Taxonomy. It seems that the forward-looking and standardised taxonomy-based metrics offer a promising path toward more transparent and comparable sustainability performance metrics.

Our paper has several contributions. First, it contributes to academic literature criticising the ESG ratings because they are inconsistently and in-transparently compiled (Berg et al., 2022; Dimson et al., 2020; Gibson et al., 2021). Second, it also lays groundwork for further empirical analysis of real EU Taxonomy data. Current literature is very scarce and limited to forecasts and expectations of affected stakeholders (Alessi et al., 2024; Hoepner and Schneider, 2022; Hummel and Bauernhofer, 2024; Seidel et al., 2024). While the regulation is not fully rolled out and firms do not fully report their sustainable activities yet, this paper presents early evidence and serves as a building stone for future analyses.

2. Literature Review

The EU Taxonomy (EU Regulation 2020/852) entered into force in 2021 and firms started mandatory reporting in fiscal year 2022. It serves as a framework for firms to classify their sustainable economic activities in a way that actual measurable financial figures can be drawn from their activities. It specifies the following green metrics: green revenue, green capital expenditure (CapEx), and green operational expenditure (OpEx). In theory, this enables an unprecedented level of transparency and comparability across firms, enabling investors to make better informed decisions about sustainable investments which in turn increases firms' transition.

Before the EU Taxonomy, it was difficult to categorise a firm or its operations and investments as sustainable. For a quantitative approach, investors as well as researchers often used ESG ratings provided by rating agencies. However, in recent studies, these ratings were heavily challenged because of the vastly differing rating criteria, scopes and weightings by the agencies that provide them (e.g. Berg et al., 2022; Dimson et al., 2020; Gibson et al., 2021). These authors conclude that different providers' ratings for the same firm diverge strongly, which raises questions about the validity and credibility of ESG ratings in the first place. Our paper adds to this conjecture by analysing the association between ESG ratings and the green KPIs reported according to the EU Taxonomy.

The field of empirical research on the EU Taxonomy is limited and based on vendors' forecasts since the regulation is still rolling out. Data from voluntary firm reporting is available for FY 2021 and mandatory reporting started in FY 2022. According to a broad analysis of seven vendors by Hoepner and Schneider (2022), mean revenue eligibility was just over 20%, while mean alignment was in the low single digits. The authors also find great variance between vendors with correlations ranging from relatively high to almost non-existent. For 2024, PwC found values of 30% revenue eligibility and 9% revenue alignment across all industries in the sample of their yearly EU Taxonomy report, suggesting an increase in firms' sustainable activities (Seidel et al., 2024). Studies on real effects of forecasted taxonomy data show that taxonomy-alignment premiums are paid to more sustainable firms as excess monthly returns (Bassen et al., 2022). This suggests that investors start implementing taxonomy metrics to assess firms' sustainability. Sautner et al. (2022), find that firms that report higher taxonomy-eligible activities paid lower interest rates for loans. However, a major shortcoming of these papers is that they use only taxonomy-aligned revenue forecasts by FTSE Russel and S&P Global respectively. Therefore, to our best knowledge, we provide the first study that empirically analyses all three green metrics and uses actual firm data. We follow Alessi et al. (2024), who develop three green KPI's based on the green taxonomy metrics reported by firms. Specifically, we first calculate a firm's Green Profit, Green Long-Termism and Speed of Transition (Alessi et al., 2024) and then combine these into the Green Score.

While researchers and policy makers advocate for the EU Taxonomy as a helpful tool to direct sustainable investments, firms suffer from regulatory uncertainty, lack of knowledge and struggle to assign their resources in such a way that their reports satisfy the authorities. Based on legitimacy theory, Hummel and Bauernhofer (2024), propose that the current state of EU Taxonomy reporting can neither be categorised into substantive nor symbolic reporting but rather as reporting in between – "endeavour to comply". This highlights the challenges that firms are facing during the taxonomy implementation despite their willingness to comply. We explore to what extent this state can be seen in the data during the first years of implementation. Similarly, O'Reilly et al. (2024) find that SMEs are willing to adapt the EU Taxonomy but they lack the required infrastructure to capture the data and outsourcing is too costly.

In sum, the taxonomy literature is still in its infancy and there are only a few empirical papers based on vendor data. On the one hand, there is an opportunity that the EU Taxonomy can relocate investments into more green projects and the reported metrics can serve as a reliable tool to measure the sustainability of firms. On the other hand, there is a lot of regulatory uncertainty and criticism is emerging on the mechanisms of the regulation itself.

3. Theoretical Background

3.1. EU Taxonomy Regulation

In recent years, the EU has worked extensively on establishing regulations, directives and guidelines for EU firms, to meet their climate and energy targets laid out by the European Green Deal in 2020. The first major development in this direction was the non-financial reporting directive (hereafter NFRD) (Directive 2014/95/EU) agreed upon in 2014 with reporting starting in FY 2017. With the NFRD, large public-interest entities were required to disclose non-financial information in their annual reports. Information included environmental, social and governance (ESG) matters. However, the scope of this directive was limited, and it lacked the necessary standardisation to be a transparent and unambiguous tool. This is why it was reformed in 2020 and replaced by the Corporate Sustainability Reporting Directive (hereafter CSRD) (Directive (EU) 2022/2464). The scope of firms was greatly increased and detailed reporting standards were introduced to allow for more standardised reports. In 2018, the EC set up a technical expert group for sustainable finance with the goal to direct investments towards a more sustainable future. Consequently, the EU Taxonomy Regulation emerged (Regulation (EU) 2020/852) and entered into force in 2020. It supports the CSRD by providing a clear classification tool for sustainable activities within a firm, specifying green metrics to be reported and thus offering a standardised numerical approach to measuring sustainability. It also aims at protecting against greenwashing practices and helps in accelerating the financing of already sustainable projects.

The EU Taxonomy defines six objectives with which a firm's activities should be aligned when determining whether they are sustainable. These objectives are: (1) climate change mitigation; (2) climate change adaptation; (3) sustainable use and protection of water and marine resources; (4) transition to a circular economy; (5) pollution prevention and control; (6) protection and restoration of biodiversity and ecosystems. It further provides an extensive catalogue describing what types of activities may be deemed sustainable. After identifying these activities, a firm is expected to calculate green metrics based on the cash flows related to them (EC, 2023): green revenue (or turnover), green CapEx and green OpEx,, and a comparison to their corresponding totals. Green revenue is defined as the proportion of the firm's net revenue that is derived from products or services that are taxonomy-aligned. It gives a static view of the contribution to the environmental goals. Green CapEx is the proportion of capital expenditures of activities that are taxonomy-aligned or part of a credible medium-term plan to achieve environmental sustainability. It offers a dynamic and forward-looking view of the firm's plans to transform their activities. Finally, green OpEx represents the proportion of operational expenditure related to taxonomy-aligned activities or to the capital expenditure plan. It contains non-capitalised costs related to maintenance and servicing of a firm's assets that are necessary to maintain effective use.

Further, a distinction must be made between taxonomy-eligible and taxonomy-aligned activities accompanied by the respective metrics. If an identified activity is listed in the catalogue, it can be deemed as taxonomy-eligible. A taxonomy-eligible activity is not sustainable per-se but rather simply any activity that is covered by the regulation's scope. To be considered taxonomy-aligned (and therefore sustainable), the firm must examine each activity in a multi-step process (see Figure 1). First, the firm screens all their activities to determine which ones fall under the scope of the EU taxonomy. This way all taxonomy-eligible activities are identified and green metrics can be calculated for these activities. The firm must then verify (1) whether an activity substantially contributes to at least one of the six previously outlined objectives; (2) whether it does not significant harm to any of the other objectives; and (3) whether it complies with the minimum social safeguards, such as the UN Guiding Principles for Business and Human Rights. If all three aspects apply, a taxonomy-eligible activity may be called taxonomy-aligned and a firm is required to calculate the green metric again, specifically for the proportion of taxonomy-aligned activities (Seidel et al., 2024).

[Insert Figure 1 about here]

The EU Taxonomy has not fully taken effect at the time of writing but is implemented stepwise. At the time of writing, data is only available up until FY 2023. Therefore, in our study, the regulation is limited in (1) the scope of firms that must report and (2) the objectives that must be considered. First, currently, only publicly listed firms with more than 500 employees are subject to the regulation. In the future, the CSRD will require more firms to report taxonomy metrics. Second, firms are currently only required to consider two out of the six environmental objectives. When identifying eligible and aligned activities, they only consider climate change mitigation and climate change adaptation. All six objectives will be required in later years. These are inherent limitations of our early evidence paper: a trade-off between an early-stage analysis and more comprehensive data where firms report their full sustainability efforts.

Nevertheless, regulators worldwide are carefully observing the effects of the EU Taxonomy. If successful, it is expected to serve as a benchmark for future green regulations in other parts of the world. Furthermore, the EU Taxonomy has global implications, since not just firms within the EU will be required to report, but also foreign firms that have significant parts of their operations in EU countries need to familiarise themselves with the regulation through their EU subsidiaries (Krämer, 2022).

3.2. ESG Ratings

ESG ratings serve the purpose of providing aggregated information about the environmental (E), social (S) and governance (G) performance of a firm. Since the EU Taxonomy relates to environmental goals, we are interested in total ESG rating and E rating. ESG ratings are prepared by third-party providers who go through a lengthy analysis process to determine a rating for a firm. They generally follow the investor-pays model, meaning the ratings are prepared independently and then to gain access, stakeholders must pay the provider for the information.

The goal of ESG ratings is to aid investors in making an informed decision for their sustainable investment strategies, but also to create a comparable metric, where the firms themselves can compare how they perform on ESG matters relative to their peers (EC, 2024). To meet this goal, providers process numerous publicly available information sources to ensure comprehensive coverage. This information includes annual reports, firm websites, NGO websites, stock exchange filings, corporate social responsibility (hereafter CSR) reports, news sources, but also personal interviews with firm representatives. Due to the large amount of data processed, providers also rely on algorithmic models and AI to improve the efficiency of their evaluation.

Methodologies for ESG ratings vary across providers, leading to significant discrepancies in ratings. As an example, LSEG breaks ESG down into these subcategories. The E pillar is comprised of resource use, emissions and innovation. The S pillar consists of workforce, human rights, community and product responsibility. The G pillar includes management, shareholders and CSR strategy. Each subcategory is assigned a relative weight based on its perceived materiality. Analysts systematically classify and interpret data from the aforementioned sources, integrating both qualitative and quantitative metrics to derive an overall ESG rating that reflects the firm's sustainability performance.

4. Hypothesis Development

In the context of measuring sustainability performance of a firm, stakeholders often use ESG ratings, as they are an easily accessible, low effort tool that can quantify, among other things, the sustainability efforts of a firm via the ESG rating. The rating that the firm receives enables easy comparability among industry peers and provides a convenient way to assess the economic situation. ESG ratings have become an essential tool for sustainable investing and their relevance has been growing alongside the growing interest in sustainable investing.

In recent academic publications however, ESG ratings were criticized for their divergence across providers and lack of transparency. As explained by Berg et al. (2022), ESG ratings, unlike credit ratings, are a concept based on values that are diverse and evolving. It is thus not surprising that the ratings diverge quite strongly between different providers, as they are competing in the market to convince investors that their way of measuring these concepts is the most accurate. They further propose the so-called "rater effect" as part of the cause for ESG inaccuracies. Borrowed from the well-known halo effect, they argue that because evaluating this type of performance requires some degree of judgment, ESG ratings are prone to this effect, where the performance of a firm in one category influences the perceived performance in other categories and find empirical evidence for it. Berg et al. (2022) conclude that divergence due to weighting and scope is to some extent desirable as there are also diverging investor preferences on scope and weight. However, most of the divergence stems from differences in the fundamental measurement approach, which is problematic, considering ESG ratings should ultimately provide objective measures. Dimson et al. (2020) provide some examples of inconsistencies in the benchmarks: "for example, Sustainalytics compares firms to constituents of a broad market index, whereas S&P compares firms to industry peers" (p. 5). Another aspect is missing data and how it is handled by different vendors. Lastly, the sheer amount of ever-expanding public information makes an objective, unified comparison almost impossible as reasons for disagreement among vendors rise.

As an alternative measurement of firm sustainability, one could use the EU Taxonomy metrics and KPIs based on those. At face value, we can argue that these measurements observe the same thing, namely sustainability efforts of a firm. However, given the inconsistencies in ESG ratings described above, we to propose the following hypothesis in alternative form:

H: Taxonomy-based green metrics are weakly associated with E(SG) ratings.

For taxonomy-based metrics, we use the Green Score which we develop from green KPIs proposed by Alessi et al. (2024). We calculate green KPIs and the Green Score using hand-collected EU Taxonomy data reported by firms in the years 2022-2023. Details of the Green Score are in the analysis section of this paper.

There are some key benefits of the EU Taxonomy data versus ESG ratings that not only make it a different measurement of sustainability, but also a more suitable one for the transparency needs of regulators and investors. The EU Taxonomy sets clear definitions for what qualifies as a sustainable activity. These definitions and following calculation rules apply to all firms, therefore the regulation enables direct comparability between firms within and across industries. Since the sustainability metrics that firms are required to report are numerical and verifiable, this financial data-based approach is less subjective than qualitative ESG ratings. Furthermore, the metrics allow for the generation of forward-looking KPIs, such as Green Long-Termism and Speed of Transition – aspects especially important in future-oriented sustainability reporting. Before the EU Taxonomy, investors would have to rely on the non-financial reporting and outlooks made by a firm, leaving room for interpretation and potential greenwashing.

We thus argue that the Green Score and green KPIs are weakly associated with E(SG) ratings. This implies that ESG ratings do not represent the objective, numerical sustainability efforts of a firm and would thus support the recent literature criticizing ESG ratings because of their non-transparency and missing objectivity.

5. Data & Methodology

For our empirical analysis, we hand-collected taxonomy data from the annual reports of European firms for the years 2022 and 2023. We start with 40 largest publicly listed firms in France, Germany and Italy according to the countries' stock indices (i.e., DAX for Germany, CAC40 for France, FTSE MIB for Italy as of December 2024). These countries were selected because they have the largest GDP in the Euro area and make up about 53% of the EUs total GDP (Statistisches Bundesamt, 2025), therefore being a representative sample. Since financial firms follow their own taxonomy regulations, they are currently not comparable to non-financial firms and were excluded from the sample.

Further, duplicates were removed for firms that show up in multiple indices (e.g. Airbus, Stellantis). This left us with a final sample of 91 firms or 182 firm-year observations. FY 2021, is the first year of the EU Taxonomy implementation and is excluded, because reporting for taxonomy-aligned activities was voluntary, and firms were still adjusting to the new reporting requirements.

The ESG ratings were collected from the EIKON database which contains ratings provided by LSEG Data & Analytics. LSEG uses a rating system that ranges from D- (the worst possible score) to A+ (the best possible score). The ratings are provided individually for E, S and G, and as the total ESG rating. For the analysis, we transformed all ratings into equally incrementing numerical values between zero and one. The outcome can be seen in table 1 below. This way, D- takes on the value zero, or the worst possible score, and A+ takes on the value one, or a perfect score, while a grade such as B would be equal to 0.637.

[Insert Table 1 about here]

We specifically extracted the E ratings and the total ESG ratings. However, many of our firms only had ratings up until FY 2022 available at the time of writing. Therefore, we always extracted the last 3 available ratings and calculated the average for each firm.

For the regression analysis we collected additional firm data from annual or sustainability reports that may influence the Green Score. We collected the total employee count, total assets and the net income of our sample firms. Another important indicator, that is directly related to sustainability, is carbon dioxide equivalent (hereafter CO_2e) emissions. Equivalent means that all greenhouse gases (hereafter GHGs) a firm emits are translated into CO_2, to create a comparable and simplified metric among all firms.

Firms are required to separate the GHG emissions into three scopes, depending on the immediacy of where they occur. Teske, 2022, describes them as follows: Scope 1 is for direct emissions from owned or controlled resources. Scope 2 is for indirect emissions from the generation of purchased energy. Scope 3 is the broadest one and includes all indirect emissions that occur in the value chain of the reporting firm (upstream and downstream). We separated the available data on GHG emissions into two variables: scope 1&2 and scope 3. This is because scope 3 is by far the most abstract, with ongoing challenges on how to measure it fully. We created a separate variable for scope 3 emissions.

Finally, one last regarding the collection of GHG emission data is that many firms report two different numbers for their emissions by scope: location-based emissions and market-based emissions and the numbers can sometimes differ strongly. Location-based reporting uses the average emission intensity of the power grid(s) that a firm is physically connected to, regardless of any contracts. Market-based reporting on the other hand reflects emissions from the purchased electricity and energy contracts are considered (Richardson, 2025). Since the market-based approach rewards firms for making smart energy procurement choices, we believe it is more in line with the future- and investment-oriented Green Score, and therefore decided to always select this approach, whenever it was available.

6. Empirical Findings

6.1. Anecdotal Evidence

This section describes isolated observations or subjective experiences encountered during gathering of taxonomy data. While they lack the statistical validity and generalisability, these important points illustrate the challenges of compliance with sustainability regulations.

An important aspect of EU Taxonomy reporting is the way firms present the taxonomy tables in their reports. During the manual data extraction process, numerous inconsistencies and errors were identified, particularly concerning the mislabelling of tables. For instance, OpEx was sometimes labelled as CapEx, or all three tables were incorrectly labelled as revenue. While such errors can often be detected through logical reasoning, inaccuracies in numerical calculations – such as incorrect sums or misplaced decimal points – pose greater challenges.

To ensure data accuracy, we attempted to correct these mistakes in our data whenever they were noticeable. However, the prevalence of such errors raises concerns about the overall reliability of the reported figures. They may undermine investor confidence in the accuracy of firms' green metrics and their adherence to reporting requirements. Given that many of the sampled firms are large multinational corporations with significant resources, it is reasonable to expect a higher level of diligence in reviewing and verifying these disclosures.

Rather than reflecting an "endeavour to comply", as introduced by Hummel and Bauernhofer (2024), one could argue that firms may be leaning more toward symbolic compliance, potentially prioritising cost savings over ensuring the accuracy and transparency of their EU Taxonomy disclosures.

6.2. Descriptive Statistics

We begin the analysis by examining the structure of the sample and creating reference points that can act as a baseline for understanding the upcoming deeper discussion. First, we plot the relative taxonomy metrics into histograms. We show all three aligned metrics relative to the eligible ones. We see that the vast majority of reported values range between 0% and 10%. For all other 10 pp intervals, the values are rather evenly distributed. For all three metrics a slight jump can be observed in the 90% to 100% range. The very large first container can be attributed to the large number of firms reporting zero values in their taxonomy metrics.

[Insert Figure 2 about here]

[Insert Figure 3 about here]

[Insert Figure 4 about here]

We consulted annual reports of a few firms to find the reasons. For example, Adidas reported zeroes for all metrics in the observation period, besides CapEx in FY 2023. According to the firm, the "[…] core business activities – the manufacturing of textiles and footwear as well as wholesale and retail thereof – are referenced by environmental objective number 4, 'Transition to a circular economy (Adidas, 2022, p. 106).'" As explained previously, in the current stage of the taxonomy roll-out, only the first two environmental objectives are required in reporting "[…] which do not reference the main economic activities of [their] industry (Adidas, 2022, p. 106)." ArcelorMittal, also reported zeroes across the board but is less transparent in reasoning. The firm mentioned that they identified a "substantial contribution to climate change mitigation for the manufacture of steel under the technical screening criteria (ArcelorMittal, 2023, p. 59)", which they could not add to taxonomy alignment because of "some gaps", mainly related to the DNSH criterion. This does however not explain why they also reported zeroes in the eligible taxonomy metrics. We conclude that our data exhibits a trade-off between conducting an early analysis of the EU Taxonomy Regulation data and having the regulation only partly rolled out, which causes many firms to propose that their activities would fit better to objectives three to six.

Next, we calculated the average eligible and aligned revenue, CapEx and OpEx as percentage of the respective total to document the extent to which the EU Taxonomy has been applied in our sample. As shown in Table 2, the proportions range from 0 to 100%. On average, eligible revenue accounts for 31.43% of the firm's total revenue. For eligible CapEx it is 46.24% and for eligible OpEx it is 34.66%. Looking at the alignment metrics, we see 8.89% for revenue, 17.89% for CapEx and 14.61% for OpEx. Compared to the PwC study for FY2023 (Seidel et al., 2024) the averages are higher. This could be attributed to our sample containing larger firms that report taxonomy data. Therefore, they may have more resources available to direct their activities to a sustainable future. As for the CapEx averages being the highest, they are in line with the findings in the study.

[Insert Table 2 about here]

We also observe a noticeable improvement of multiple percentage points in all metrics year over year. The largest improvement can be seen in eligible revenue, which goes up by 9.87 pp, while the smallest increase can be observed for aligned OpEx with 1.93 pp increase. Overall, these are satisfactory results that hopefully provide a glimpse of further eligibility and alignment improvements in the years to come.

Table 3 reports ESG ratings and we see that the average E rating is the same as the average ESG rating: firms have an average rating A-. According to the LSEG, all firms with ratings between A+ and A- are titled ESG leaders because their ratings indicate excellent relative ESG performance and a high level of transparency (LSEG, 2024).

[Insert Table 3 about here]

We should mention a limitation of the paper, namely the measurement process of ESG ratings. We had to come up with a numerical scale that attributes one specific value to a given rating, but LSEG uses numerical ranges behind their ratings. So, while our A- rating stands for exactly 0.819, it can in reality range between 0.75 and 0.83 according to the LSEG scale. The true value behind a rating is unobservable, since LSEG only publishes the final ratings. This can therefore lead to inaccuracies in our ESG data, because they are limited to an estimation.

6.3. Green Score

Our goal is to measure sustainability performance of large EU firms based on the three green KPIs proposed by Alessi et al. (2024): Green Profit, Speed of Climate Transition and Green Long-Termism. These KPIs will show how green a firm currently is and also how seriously it considers sustainability for its future operations. They are calculated as follows:

$$Green\ Profit\ = green\ revenue - green\ OpEx$$

$$Speed\ of\ Transition = \ln(\frac{green\ CapEx}{green\ revenue})$$

$$Green\ Long - Termism = \ln(\frac{green\ CapEx}{green\ OpEx})$$

Green Profit is a static KPI that can be used to identify firms with the best green margins. Speed of Transition focuses on the transition towards a low carbon economy. The natural logarithm helps in norming the results around zero and makes interpretation easy – firms with transition speed over zero embrace climate transition, while the ones with negative results are reluctant to it. Finally, Green Long-Termism measures how focused a firm is in investing in new green assets. Here again, negative results signal a focus on existing green assets and a reluctance to invest in new green assets.

Before calculating the KPIs, we should mention a shortcoming of the suggested formulas. Alessi et al. (2024) do not address the weight of the taxonomy data. We illustrate what is meant by this with an example. Assume a car manufacturer that emits large amounts of CO_2 during the production process and their taxonomy alignment metrics are all extremely low relative to their total revenue, CapEx and OpEx. An analyst would probably argue that this is not a sustainable firm. Yet, applying the proposed metrics, they could get one of the best Green Long-Termism scores simply by their aligned CapEx being a high multiple of their aligned OpEx, despite the absolute numbers being almost negligible.

To improve the Speed of Transition and Green Long-Termism, the original formula should be multiplied by the total aligned values divided by the total eligible values. This ensures that measures of the proposed KPIs consider the extent of a firm's effort. Here we assume that taxonomy eligibility is comprised of all business activities that the firm could possibly transform into taxonomy-aligned. Consequently, the higher the firm's taxonomy alignment relative to taxonomy eligibility, the more sustainable the firm is. The formulas then become:

$$Green\ Profit = aligned\ revenue - aligned\ OpEx$$

$$Speed\ of\ Transition = \ln\left(\frac{aligned\ CapEx}{aligned\ revenue}\right) \times \left(\frac{aligned\ CapEx + aligned\ revenue}{eligible\ CapEx + eligible\ revenue}\right)$$

$$Green\ Long-Termism = \ln\left(\frac{aligned\ CapEx}{aligned\ OpEx}\right) \times \left(\frac{aligned\ CapEx + aligned\ OpEx}{eligible\ CapEx + eligible\ OpEx}\right)$$

Next, we calculate a firm's Green Scores based on these three KPIs and investigate its correlation to ESG rating We standardise Green Profit, Speed of Transition and Green Long-Termism with help of the Z-Score:

$$z_x = \frac{x - \bar{x}}{\sigma_x}$$

x is the respective KPI for a firm, $\bar{x}$ is the mean KPI across all firms and σ_x is the standard deviation. This way, each of the KPIs receive a standardised z-score with the help of which a Green Score is calculated, which weighs all three KPIs equally.

$$Green\ Score = z_{GreenProfit} \times \frac{1}{3} + z_{Long-Termism} \times \frac{1}{3} + z_{Transition\ Speed} \times \frac{1}{3}$$

If x follows a standard normal distribution, the z-score ranges between zero and one and so would the Green Score. However, the true values of Green Score range between -1.9953 and 2.2639. This suggests that to no surprise, our KPIs do not follow a standard normal distribution. While it is difficult to define a minimum or maximum for the Green Score based on this information, we provide estimation ranges. Consulting the three KPIs, we see that Green Profit ranges from -0.5111 to 4.9906. Speed of Transition ranges from -4.1165 to 1.2071 and Green Long-Termism between -2.1105 and 3.8950. Because these values do not differ too much from the range of the Green Score, we conclude that theoretically a minimum of -3 and a maximum of +3 can be achieved, with 3 essentially being a perfect score.

[Insert Figure 5 about here]

From the distribution of the Green Scores shown in Figure 5, we see that the vast majority of values between -1 and 1. We consider Green Score of less than -1 especially bad (exhibited by 3% of observations), while a score that is better than 1 means that a firm is handling sustainability exceptionally well (exhibited by 8% of observations). The average Green Score in our sample is 0.0142, with 27% of observations having a score higher than zero.

Investigating the best performers according to the Green Score, we see that six out of the top 10 firms are energy providers and three are car manufacturers. These firms are among the largest sample firms in terms of revenue. The positions seem reasonable because energy providers usually already have part of their operations aimed at sustainability (e.g. producing and providing green energy). Thus, it is easier for them to identify taxonomy-aligned revenues and investments, and they are naturally a bit ahead of other firms. Further, we posit that larger firms have more available income to invest in sustainable activities. It is likely easier for them to adapt to new regulation requirements as they can allocate resources to expert staff and infrastructure. On the other hand, the worst performers according to the Green Score are more diverse. They are mostly industrial and manufacturing firms, operating in areas such as oilfield services, steel manufacturing or manufacturing with rubber-based materials. These firms traditionally operate in emission intensive industries, which makes taxonomy alignment a challenge.

However, the data has some anomalies. For the ESG ratings, a lot of firms usually get the same ranking in the distribution because there is a limited number of ratings they can receive. For the green metrics, there tends to be a lot of firms in the middle of the distribution, where Green Long-Termism or Transition Speed are zero (because firms reported zero values). In our sample, a lot of firms received a Green Score of -0.1664 because they reported all zero values. Specifically, there are 74 out of 182 observations like this.

6.4. Green Score versus ESG Score

[Insert Figure 6 about here]

Figure 6 shows the Spearman's rank correlation between the E rating and the Green Score for our sample. It visualises the previously described anomaly (many firms with Green Score of -0.1664). For the E ratings, the data points jump in fixed intervals because there are only 12 potential ratings possible. The figure suggests no clear correlation between the two rankings.

No specific form of a relation can be visualised; rather, the points are scattered across the entire field. This conjecture is confirmed by Pearson's correlation coefficients between the Green Score and the E(SG) rankings presented in Table 4.

[Insert Table 4 about here]

Presented correlation coefficients are close to zero and statistically insignificant. The strongest positive correlation, albeit insignificant, is observed between the Green Score and the E rating (0.1106). This suggests that there are at least some commonalities between the two variables, albeit little. The ESG rating has an even lower correlation with the Green Score (0.0846). These findings at best indicate only weak association between the Green Score and E(SG) ratings as hypothesised.

In Table 5, we look at the correlation coefficients between the disaggregated Green Score (i.e., Green Profit, Speed of Transition, Green Long-Termism) and E(SG) ratings. All correlations are statistically insignificant for Green Long-Termism and Speed of Transition. On the other hand, correlation coefficient between Green Profit and the E rating is highly statistically significant. The coefficient value of 0.2072 suggests a weak association as hypothesised. The ESG rating is not statistically significantly correlated with Green Profit.

[Insert Table 5 about here]

We argue that Green Profit is the only green KPI that is not strictly forward looking and does not use green CapEx in its calculation. It rather shows a static view of the firm's current green performance. Therefore, it seems reasonable that the E(SG) ratings are only related to this metric, as it is the most similar to them in its measurement.

Poor Reporting

As explained in section 6.1, several firms that made some mistakes when reporting the taxonomy tables. Most commonly, we assume that the template for the green revenue table was used for all metrics without changing the table in the relevant areas. This results in eligible and aligned CapEx and OpEx being incorrectly labelled revenue or turnover. In rarer cases, percentages or decimals would be missing, or the numbers would not add up altogether.

We further argue that these firms did not apply the expected level of diligence in handling this new regulation, potentially reflecting a symbolic rather than substantive commitment to sustainability disclosure. This aligns with legitimacy theory, which suggests that some firms engage in symbolic reporting – appearing compliant without ensuring the accuracy or transparency of their disclosures. If this is the case, it is reasonable to assume that these firms may also struggle to accurately report their taxonomy-based metrics.

Consequently, and as a further robustness test, we classify firms as poor-quality reporters if they showed notable reporting mistakes. We create a dummy variable *poor_reporting* that takes the value one if the firm has made mistakes in the reporting of their taxonomy metrics and zero otherwise. We removed poor-quality reporters for the sample (34 observations) and tested the correlations again. Any meaningful increase in correlations would signal that there is indeed a relationship between the scores, however it was obfuscated by poor reporting. Table 6 reports the results.

[Insert Table 6 about here]

As opposed to the full sample, for high-quality reporters we find that Green Score is significantly positively correlated with E rating (0.1106). Similar to the full sample, we document that Green Profit is significantly correlated with E rating (0.2125). We therefore conclude that poor reporting is not a very important factor for the lack of correlations between green metrics and E(SG) ratings but it has some influence.

Regression Analysis

A regression analysis shall provide more insight into the relationship between Green Score and the E(SG) rating and also show what other factors drive the Green Score.

$$Green\ Score_{it} = \beta_0 + \beta_1 \times ESG^*_{it} + \beta_2 \times \ln(Revenue_{it}) + \beta_3 \times ln(Employees_{it}) + \beta_4 \times ROA_{it} + \beta_5 \times Capital\ Intensity_{it} + \beta_6 \times Emission\ Intensity_{it} + \beta_7 \times Poor\ Reporting_{it} + \delta_c + \gamma_s + \tau_t + \varepsilon_{it}$$

The dependent variable is the Green Score and the main independent variable can be ESG rating or E rating. We test whether ESG rating is associated with taxonomy-based green metric. All variables are for firm i in year t. We included several control variables. While revenue directly controls for firm size and scale of operations, the employee count controls for labour-intensive vs. capital-intensive business. We used a log transformation to normalise the scale and prevent very large outliers from influencing the sample too much. This way we can also account for diminishing returns: revenue and employee count increases do not have linear effects but rather become less impactful with increasing size. We also include profitability variable, ROA, which is calculated by dividing the total income by the total assets. Another factor that might influence the Green Score is capital intensity. This is because the score partially depends on CapEx for the Speed of Transition and Green Long-Termism calculations. Firms with high capital intensity may be more likely to invest into new forward-looking and sustainable projects which in turn improves their Green Score. We also decided to include the variable for scope 1&2 CO_{2e} emissions. It is a static indicator of the environmental damage and will indicate whether the emissions a firm produces is associated to the Green Score. Finally, we include the *Poor Reporting* dummy again to test for the potential impact of firms that made mistakes in their EU Taxonomy reporting. We further control for country fixed-effects (δ_c), for sector fixed-effects (γ_s) and for time fixed-effects (τ_t).

We also conducted the usual tests to assess the quality of control variables and the goodness of fit of our model. We do not report them in the paper. The variance inflation factor suggests no clear signs of multicollinearity and confirms that the variables were chosen appropriately.

[Insert Table 8 about here]

The main regression results are shown in Table 8. We find that the E rating does not significantly explain variation in the Green Score. It may therefore not capture the green performance as defined in the EU Taxonomy. The employee count and poor reporting are also insignificant. However, the other control variables have statistically significant coefficients. We conclude that a 1% increase in revenue is associated with an increase of 0.3163 in Green Score. This suggests that larger firms are more taxonomy-aligned and more likely to get a better Green Score. The coefficient for ROA is 0.0086, which suggests that more profitable firms tend to be more aligned with the Green Score, possibly because they can afford more transition investments. The same goes for capital intensity. With a coefficient of 1.483 suggests quite a strong effect, that can be explained by (aligned) CapEx directly being used as a measurement in the Green Score. Lastly, the positive coefficient CO_{2e} emissions suggests that increase in emission intensity is associated with increase in the Green Score. This means that firms that emit more GHG tend to higher Green Score. However, as we found that Green Score also correlates firm size, and large firms tend to have more operations that may emit GHG, this result may indicate that many higher-emission firms are also working the hardest on reducing their impact on the environment via sustainable investments. We find that Italian and German firms have significantly higher Green Scores than French ones. Also, firms had higher average Green Scores in FY2023 compared to FY2022, suggesting a significant improvement year-over-year, albeit a small one.

We also replace the independent variable with the ESG rating but do not tabulate the results. While ESG rating is not significantly associated with Green Score, coefficients for the controls show results similar to Table 8.

6.5. Additional Analyses

Country Comparison

We selected the sample based on the EU countries with the largest GDP. In this section we compare results for the three countries. As shown in Table 10, Germany (35% of observations) has the highest absolute values of all alignment and eligibility metrics. When considering the ratios between alignment and eligibility, a different trend emerges: Germany performs worst, followed by France (39% of observations). Italy (26% of observations), as the smallest economy of the three, exhibits the best ratios, indicating a faster adaptation of the EU Taxonomy Regulation and a stronger commitment to transform their business activities towards sustainability. 54% of eligible revenues in Italy are aligned. The perhaps most forward-looking metric, aligned CapEx, indicates that about 83% of the eligible investments that are aligned, highlighting a strong future-oriented approach. Finally, the aligned OpEx of about 61% suggests that operations are also carried out predominantly green.

However, this could also be influenced by the sectoral composition of the Italian economy, as some industries, such as energy providers may naturally align more easily with the EU Taxonomy. As is shown before, energy providers generally perform best according to the Green Score because of sustainability naturally being part of their offering portfolio. Italy specifically is characterised by a low concentration of market share among its largest energy providers and therefore many of them are part of the Italian index. In comparison, Germany's ratios are 14% for revenue, 33% for CapEx and 30% for OpEx, leaving a lot of room for improvement and a large gap between the two countries. This may, however, also be partially explained by Germany's broader industrial mix, which includes more high-emission sectors that face greater regulatory and technological barriers to alignment.

[Insert Table 10 about here]

We further extended the analysis by calculating the green metrics for each country. In Table 11, we see that Germany performed best in Green Profit, which is expected, given that this metric reflects absolute green revenue. Since the German firms of our sample on average earn higher overall revenues, they have more room for parts of them to be green. Italy's Green Profit is below but relatively close to France's, despite much lower absolute revenues, again suggesting a higher focus in this area by Italy. Italy's Green Long-Termism score is a multiple of the scores by Germany and France, consistent with the findings of the taxonomy measures of before and further solidifying the aspect that Italy's strong green CapEx suggests a long-term orientation in their economy. For Speed of Transition, all three countries exhibit negative values, suggesting a slow and somewhat reluctant shift towards sustainability. However, this trend may improve as firms become more accustomed to EU Taxonomy requirements and strive to enhance their alignment metrics over time.

[Insert Table 11 about here]

Overall, these findings highlight Italy's comparatively stronger commitment to sustainability, both in terms of current alignment and forward-looking investment strategies, while Germany, despite its economic dominance, lags in relative alignment.

Removing Zeroes

A potential explanation for the lack of significant correlation between the Green Score and E(SG) rating is the presence of reported zeroes in the EU Taxonomy metrics. Due to the current regulatory environment, reported zeroes in many cases do not indicate an absence of sustainability efforts by firms. Instead, they often reflect the fact that certain firms do not yet see themselves as being required to disclose these metrics. As a result, zero values may be misleading and could negatively influence our statistical analysis.

We test the robustness of our results to excluding zero values. We removed all firms from the sample that reported zeroes across all three metrics (Green Revenue, Green CapEx, Green OpEx). This adjustment eliminated 33 firms, leaving a sample of 58 firms 116 observations), all of which reported at least some level of taxonomy alignment.

Upon repeating the correlation analysis with this subsample, we continued to find no statistically significant correlation between the Green Score and ESG rating (Table 12). But we find a weak significant correlation between Green Score and E rating as hypothesised. These findings suggest that the lack of correlation is not merely an artifact of zero-inflated data but rather indicative of a deeper misalignment between ESG ratings and taxonomy-based green metrics. If the presence of zeroes had been the primary factor distorting the results, their removal should have led to a stronger correlation, which was not the case.

[Insert Table 12 about here]

We rerun the regression analysis and show results in Table 13. For the goodness of fit, we see that the adjusted R^2 has improved. Looking at the relationship between Green Score and E rating, the coefficient stays insignificant. Interestingly, the coefficient of the control emission intensity is no longer statistically significant. Finally, the effects of revenue and capital intensity on the Green Score become substantially stronger, while the year fixed effect is now also statistically insignificant. We obtain qualitatively similar results when the main dependent variable is ESG rating (results untabulated).

[Insert Table 13 about here]

7. Discussion

From our analyses we conclude that the regression model and the correlations show only weak positive association between the Green Score/Green Profit and E rating. Thus, the results are in line with our predictions. The results also support the literature criticising ESG ratings for being non-transparent and inconsistent. We show that the EU Taxonomy in contrast provides forward-looking and standardised matrics not necessarily of how sustainable a firm currently is, but how active it is in transforming their operations into more sustainable future practices.

To this end, we found that the Green Score is positively influenced by profitability and firm size. This suggests that financially stronger firms can more readily invest into taxonomy-aligned projects or have the necessary capabilities to better comply with the regulatory requirements of the EU Taxonomy. Further, capital intensity is also relevant, which is intuitive since the Green Score is partially calculated using green CapEx and since new sustainable projects are often capital-intensive, firms with more capital available will have an easier time investing.

Finally, in our country-by-country analysis Italy emerges as a clear winner, its firms being the most taxonomy-eligible and taxonomy -aligned. This is also reflected in the regression results, where Italian firms on average perform significantly better than French ones, while German firms are only slightly better than French.

In sum, our analyses show that sustainability is a multi-faceted construct and it remains challenging to capture it in plain numbers and metrics. While ESG ratings have been criticised, they are still largely based on qualitative firm information which suggests the measures firms are taking to become more sustainable. We conclude that there is no "best" sustainability measurement yet, despite the EU Taxonomy metrics being a good approach. Rather than being the ultimate sustainability performance metric, it primarily looks at how strongly a firm is investing in green projects, regardless of its current practices and emissions. For this reason, if one is truly interested in the sustainability performance of a firm, one should not strictly consult the EU Taxonomy metrics but consider them in combination with the ESG ratings.

8. Conclusion

The contribution of this paper is an early, real data driven analysis of the EU Taxonomy Regulation. We highlight the practical challenges faced by the implementation as well as an outlook for its trajectory. We find that some firms struggle to accurately and correctly report taxonomy-based metrics, while other firms rely on the slow implementation and current limited scope to avoid reporting as of now.

We also introduce our own metric to measure sustainability efforts of a firm, particularly with a future orientation, the Green Score. It combines green KPIs calculated from reported taxonomy data by firms: Green Profit, Green Long-Termism and Speed of Transition.

Finally, we contribute to the literature criticising ESG ratings by comparing the E(SG) rating to the Green Score. We find no statistically significant correlation between the Green Score and ESG rating, but we find that the Green Score and Green Profit are weakly associated with E rating. Our regression analysis shows that the two scores are not associated. This further suggests that ESG ratings may lack the objectivity transparent measuring practices, while taxonomy-based metrics use a more objective and comparable numerical approach to sustainability. We suggest that the Green Score could be considered an alternative for ESG ratings because it is based on objective metrics reported by the firms and takes a forward-looking perspective. However, the Green Score's focus on investments and future sustainability also limits it power to provide the 'ultimate' sustainability performance metric.

Our study has a few limitations. The EU Taxonomy data collected stems from a very early, not fully rolled-out version of the regulation and not all sustainability objectives are yet mandatory. One should repeat this analysis once the regulation is in full effect. Also, despite the representative firm sample of the EU economy, many countries were not considered. It would be particularly important to investigate how weaker economies and smaller firms handle the regulation, which is another avenue for future research.

In sum, our study highlights the importance of robust, transparent, and standardised sustainability metrics. As the EU Taxonomy continues to be implemented, policymakers, investors, and firms alike must consider how to improve reporting accuracy, enhance comparability, and ensure that sustainability assessments align with real-world impact. Our findings not only contribute to the academic literature on sustainability measurement but also offer practical implications for the future of sustainable finance and corporate accountability.

References

Alessi, L., Cojoianu, T., Hoepner, A. G., & Michelon, G. (2024). Accounting for the EU Green Taxonomy: exploring its concept, data and analytics. *Accounting Forum*, 365-371.

Bassen, A., Kordsachia, O., Lopatta, K., & Tan, W. (2022). Revenue Alignment with the EU Taxonomy Regulation in Developed Markets. *SSRN*, 16-25.

Berg, F., Kölbel, J., & Rigobon, R. (2022). Aggregate Confusion: The Divergence of ESG Ratings. *Zurich Open Repository and Archive*.

Bongaerts, D. C. (2012). Tiebreaker: certification and multiple credit ratings. *Journal of Finance*, 113-152.

Brander, M., & Davis, G. (2012). Greenhouse Gases, CO2, CO2e, and Carbon: What Do All These Terms Mean? *Econometrica*, 1-3.

Deutsche Telekom. (2021). Interim report Q1 2021. https://report.telekom.com/interim-report-q1-2021/.

Deutsche Telekom. (2023). Annual report 2023. https://report.telekom.com/annual-report-2023/management-report/.

Dimson, E., Marsh, P., & Staunton, M. (2020). Divergent ESG ratings. *SSRN*.

Downie, J., & Stubbs, W. (2011). Evaluation of Australian companies' scope 3 greenhouse gas emissions assessments. *Journal of Cleaner Production*, 161-162.

European Commission. (2023). *EU taxonomy for sustainable activities*. Retrieved April 17, 2025, from https://finance.ec.europa.eu/sustainable-finance/tools-and-standards/eu-taxonomy-sustainable-activities_en

European Commission. (2024). *EU Taxonomy Navigator*. Retrieved April 17, 2025, from https://ec.europa.eu/sustainable-finance-taxonomy/

European Parliament and the Council . (2014). Directive 2014/95/EU of the European Parliament and of the Council. *Official Journal of the European Union*.

European Parliament and the Council. (2020). Regulation (EU) 2020/852 of the European Parliament and the Council. *Official Journal of the European Union*.

European Parliament and the Council. (2022). Directive (EU) 2022/2464 of the European Parliament and of the Council. *Official Journal of the European Union*.

Gibson, B., Krueger, P., & Schmidt, P. S. (2021). ESG Rating Disagreement and Stock Returns. *Financial Analysts Journal*.

Gujarati, D. N., & Porter, D. C. (2009). *Basic Econometrics.* New York, NY: McGraw-Hill Irwin.

Hoepner, A. G., & Schneider, F. I. (2022). EU Green Taxonomy Data - A First Vendor Survey. *De Gruyter*, 229-237.

Hummel, K., & Bauernhofer, K. (2024). Consequences of Sustainability Reporting Mandates: Early Evidence on the EU Taxonomy Regulation. *SSRN*, 20-23, 37-41 .

Krämer, N. (2022, April 28). *The Global Impact of EU Taxonomy on non-EU Green Taxonomies*. Retrieved from Greenomy: https://www.greenomy.io/de/blog/eu-taxonomy-global-impact

LSEG. (2024). Environmental, Social and Governance scores from LSEG. *LSEG Data & Analytics*.

O'Reilly, S., Gorman, L., Mac An Bhaird, C., & Brennan, N. M. (2024). Implementing the European Union Green Taxonomy: implications for small- and medium-sized enterprises. *Accounting Forum*, 419-422.

Reuters. (2021). Mediaset's top investor ups stake in peace deal with Vivendi. *Reuters*, https://www.reuters.com/business/media-telecom/vivendi-mediaset-fininvest-agree-end-dispute-2021-07-22/.

Richardson, B. (2025). *Zevero*. Retrieved from https://www.zevero.earth/blog/location-vs-market-based-carbon-reporting

Sautner, Z., Yu, J., Zhong, R., & Zhou, X. (2024). The EU Taxonomy and the Syndicated Loan Market. *SSRN*, 11-14.

Seidel, B., Schellhas, C., Gehrke, N., McClellan, A., Ladusch, A., Kolm, S., & Böcker, S. (2024). *EU Taxonomy Reporting 2024 - Analysis of the financial and non-financial sector*. PwC.

Statistisches Bundesamt. (2025). *destatis*. Retrieved April 19, 2025, from https://www.destatis.de/Europa/EN/Topic/Key-indicators/EconomyFinance.html

Teske, S. (2022). *Achieving the Paris Climate Agreement Goals*. Institute for Sustainable Futures, Sydney: Springer.

Figures

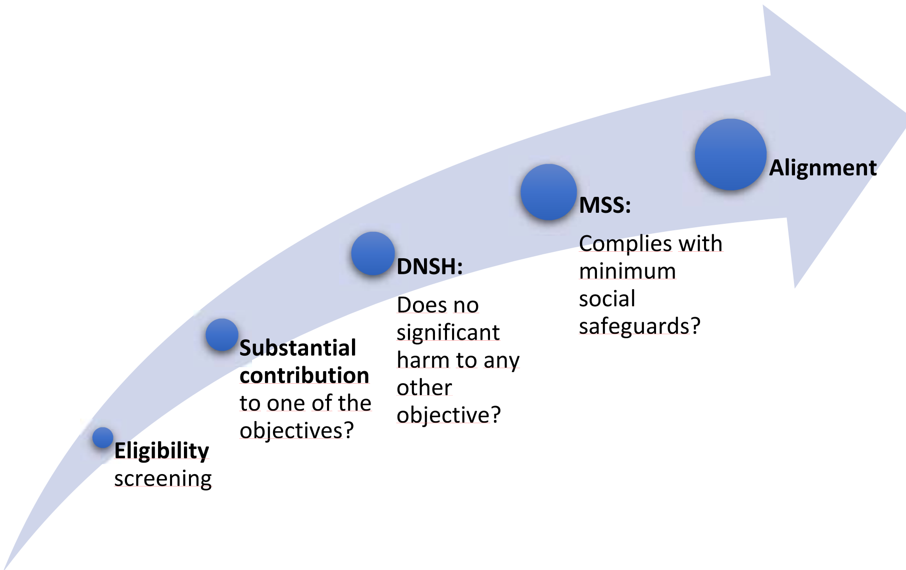

Figure 1: Steps taken for an activity to become aligned

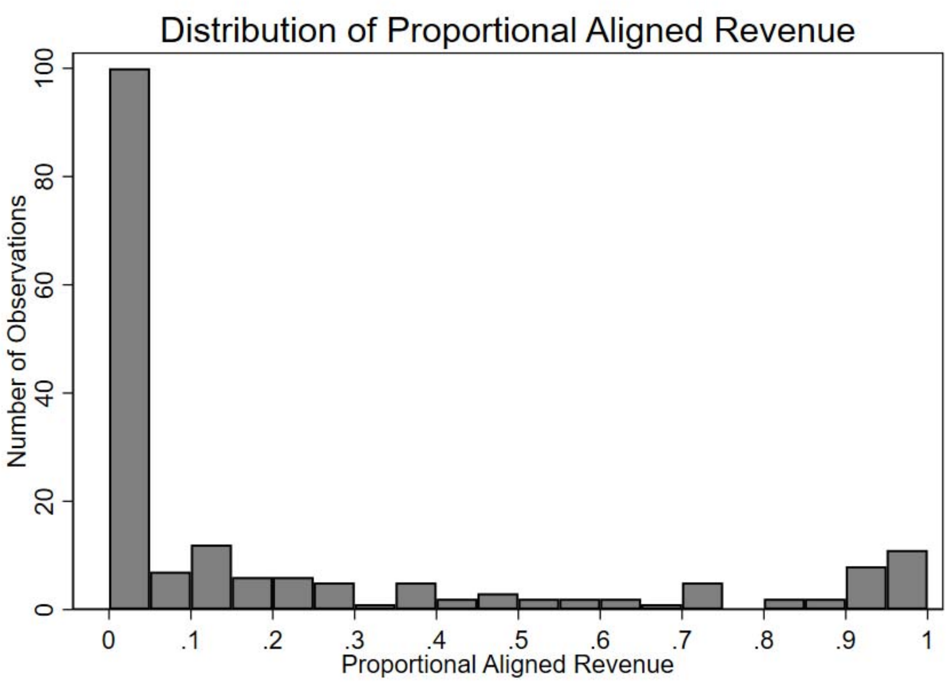

Figure 2: Distribution of proportional aligned revenue.

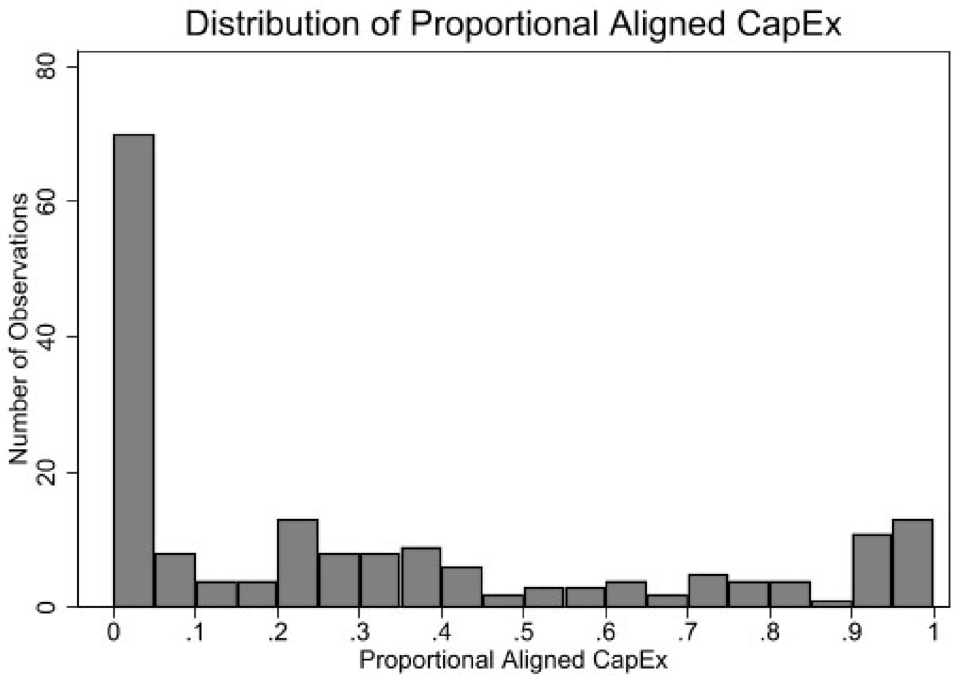

Figure 3: Distribution of proportional aligned CapEx.

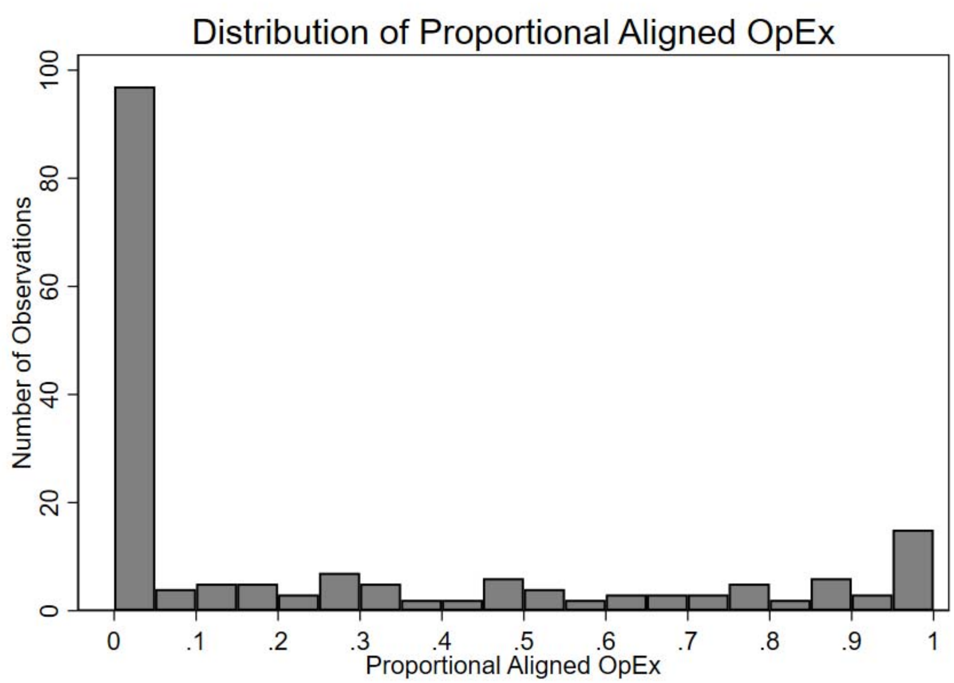

Figure 4: Distribution of proportional aligned OpEx.

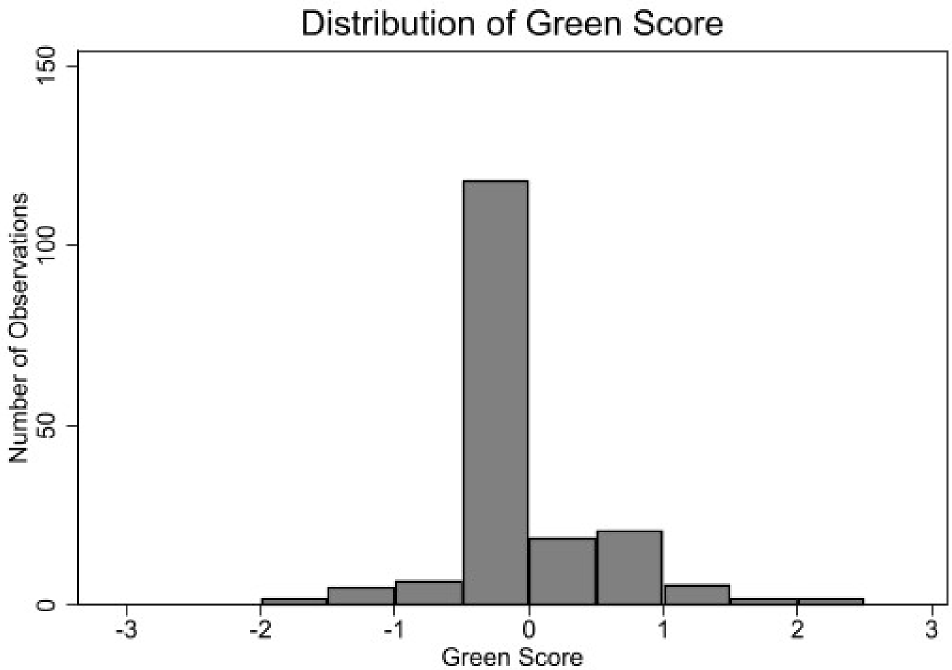

Figure 5: Distribution of Green Score.

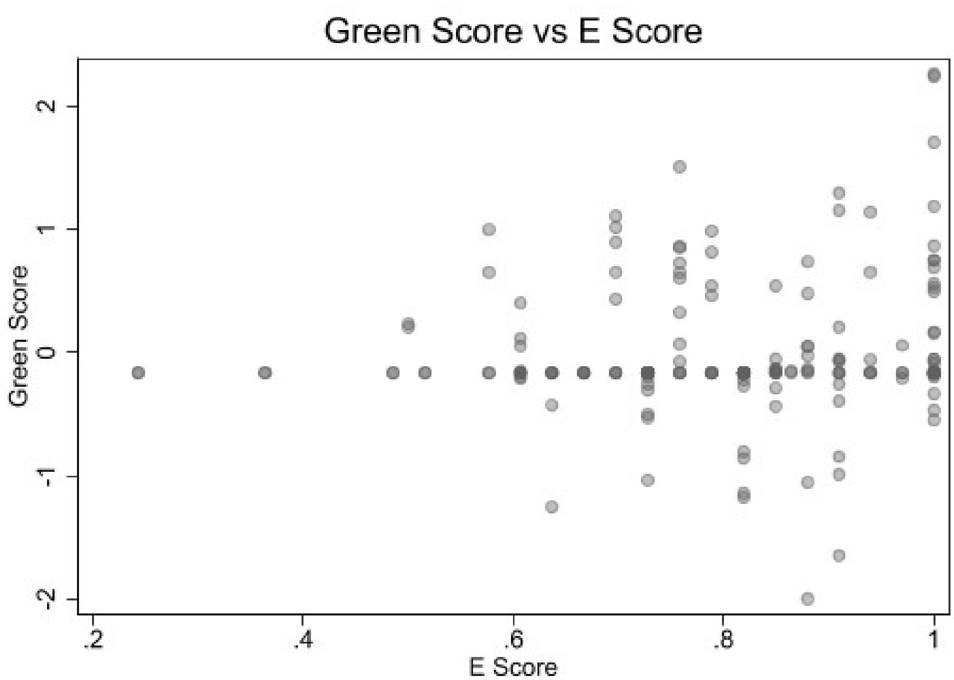

Figure 6: Green Score vs. E Rating.

Tables

Table 1: E(SG) ratings transformed into numerical values

RATING	NUMERCAL VALUE
A+	1
A	0.91
A-	0.819
B+	0.728
B	0.637
B-	0.546
C+	0.455
C	0.364
C-	0.273
D+	0.182
D	0.091
D-	0

Table 2: Average sampled EU Taxonomy metrics (as percentage of firm totals)

% of firm total	**Eligible revenue**	**Eligible CapEx**	**Eligible OpEx**	**Aligned revenue**	**Aligned CapEx**	**Aligned OpEx**
Overall (N=182)	31.43%	46.24%	34.66%	8.89%	17.89%	14.61%
2022 (N=91)	26.44%	42.52%	30.97%	7.24%	16.57%	13.62%
2023 (N=91)	36.31%	49.87%	38.22%	10.49%	19.17%	15.58%

Table 3: Average ESG ratings

	E	**ESG**
Score	0.7996	0.8036
Grade	A-	A-

Table 4: Correlation coefficients between Green Score and ESG ratings

	ESG	**E**
Green Score	0.0846	0.1106
p	*(0.2588)*	*(0.1396)*

Significance levels: 0.01 (***), 0.05 (**), 0.1 (*). N=182

Table 5: Correlation coefficients between green KPIs and ESG ratings

	ESG	**E**
Green Profit	0.0999	0.2072***
p	*(0.1822)*	*(0.0053)*
Long-Termism	0.1166	0.0889
p	*(0.1190)*	*(0.2351)*
Transition Speed	-0.0736	-0.1095
p	*(0.3261)*	*(0.1434)*

Significance levels: 0.01 (***), 0.05 (**), 0.1 (*). N=182

Table 6: Correlation coefficients controlled after excluding poor reporters

	ESG	**E**
Green Score	0.0846	0.1106*
p	*(0.3744)*	*(0.0877)*
Green Profit	0.0765	0.2125***
p	*(0.3556)*	*(0.0095)*
Long-Termism	0.0296	0.0258
p	*(0.7206)*	*(0.7552)*
Transition Speed	0.0159	-0.0141
p	*(0.8476)*	*(0.8645)*

Significance levels: 0.01 (***), 0.05 (**), 0.1 (*). N=148

Table 8: Main regression results

Green Score	Coefficient	Standard error	P > |t|
E rating	**-0.4617**	**0.3249**	**0.158**
ln(revenue)	0.3163***	0.0661	<0.001
ln(employees)	0.0176	0.0304	0.562
ROA	0.0086***	0.0032	0.008
Capital intensity	1.4827***	0.4031	<0.001
Poor reporting	-0.1012	0.1417	0.476
CO_{2e} **intensity**	0.0965**	0.0392	0.015
Country (Germany)	0.1757*	0.0987	0.077
Country (Italy)	0.4845**	0.2062	0.020
Year (2023)	0.0342***	0.8555	<0.001

Significance levels: 0.01 (***), 0.05 (**), 0.1 (*). N=182

Table 10: Average aligned and eligible metrics by country

	Germany (N=64)	France (N=70)	Italy (N=48)
Aligned Revenue	3854.95	2757.61	2394.52
Aligned CapEx	1550.10	637.31	864.98
Aligned OpEx	405.78	258.60	112.54
Eligible Revenue	28230.24	12385.54	4428.15
Eligible CapEx	4639.03	1676.71	1036.02
Eligible OpEx	1356.08	574.97	185.02

Table 11: Average green KPIs by country

	Germany (N=64)	France (N=70)	Italy (N=48)
Green Profit	3395.28	2499.01	2281.99
Long-Termism	0.307	0.252	0.739
Speed of Transition	-0.200	-0.356	-0.381

Table 12: Correlation table between Green Score and ESG rating (zeroes removed)

	ESG	E
Green Score	0.0461	0.0623
p	*(0.6886)*	*(0.5882)*
Green Profit	-0.0215	0.1766*
p	*(0.8285)*	*(0.0702)*
Long-Termism	-0.0236	-0.0585
p	*(0.8272)*	*(0.5882)*
Transition Speed	0.0619	0.0015
p	*(0.5602)*	*(0.9886)*

Significance levels: 0.01 (***), 0.05 (**), 0.1 (*). N=116

Table 13: Regression results for the main independent variable (zeroes removed)

Green Score	Coefficient	Standard error	P > |t|
E rating	**-0.6029**	**0.3785**	**0.117**
ln(revenue)	0.4647***	0.0795	<0.001
ln(employees)	0.0535	0.0369	0.153
ROA	0.0083**	0.0041	0.045
Capital intensity	2.0384***	0.4304	<0.001
Poor reporting	0.0058	0.1706	0.973
CO_{2e} intensity	0.0611	0.0909	0.504
Country (Germany)	0.0886	0.1821	0.628
Country (Italy)	0.4466**	0.2199	0.047
Year (2023)	0.1154	0.1057	0.280

Significance levels: 0.01 (***), 0.05 (**), 0.1 (*). N=116