Scope 3 Emissions: Data Quality and Machine Learning Prediction Accuracy

Quyen Nguyen [1] Ivan Diaz-Rainey [1]

Adam Kitto [2] Ben McNeil [2,3]

Nicholas A. Pittman [2]

Renzhu Zhang [1]

10 August 2022

Abstract

This paper explores the quality of Scope 3 emission data in terms of divergence and composition, and the performance of machine-learning models in predicting Scope 3 emissions. We do so using the Scope 3 emission datasets of three of the largest data providers (Bloomberg, Refinitiv Eikon, and ISS). We find considerable divergence between third-party providers, making it difficult for investors to know their 'real' exposure to Scope 3 emissions. Surprisingly, divergence exists between the datasets for emissions values that have been reported by firms (68% identical data points between Bloomberg and Refinitiv Eikon). The divergence is even larger for ISS when it adjusts reported values using its proprietary models (0% identical data points). With respect to the composition of Scope 3 emissions, firms generally report incomplete compositions, yet they are reporting more categories over time. There is a persistent contrast between relevance and completeness in the composition of Scope 3 emissions across sectors, as irrelevant categories such as travel emissions are reported more frequently than relevant ones, such as the use of products and processing of sold products. We also find that the application of machine learning algorithms can improve the prediction accuracy of the aggregated Scope 3 emissions (up to 6%) and its components, especially when each category is estimated individually and aggregated into the total Scope 3 emissions values (up to 25%). It is easier to predict upstream emissions than downstream emissions. Prediction performance is primarily limited by low observations in particular categories, and predictor importance varies by category. We conclude that users of the Scope 3 emission datasets should consider data source, quality and prediction errors when using data from third-party providers in their risk analyses.

Keywords: Scope 3 emissions; Carbon footprint; Climate finance; Machine learning; transition risk; Errors-in-variables

JEL codes: C89, G17, Q51, Q54

Acknowledgments and Funding: This is research is the product of a commercialization agreement between University of Otago and EMMI (Distributed Carbon Pty Ltd) that builds on prior related research (namely Nguyen et al., 2021). This paper has benefitted from comments received during 4th EU JRC Summer School in Sustainable Finance (July, 2022), an IAEE Webinar (June, 2022), the Machine Learning for the Climate Tech Industry Virtual Conference (June, 2022) and a Tsinghua Green Finance Lecture Series (Tsinghua University, May 2022). The paper will also be presented at the 5th GRASFI Conference (University of Zurich, 2022). The usual disclaimer applies.

---

1 Climate and Energy Finance Group (CEFGroup), Department of Accountancy & Finance, Otago Business School, University of Otago | Dunedin, 9016, New Zealand

[2] EMMI (Distributed Carbon Pty Ltd) | 710 Collins St, Docklands Victoria, 3008, Australia

[3] Climate Change Research Centre, Faculty of Science, University of New South Wales, UNSW, 2052

1. Introduction

Corporate carbon footprints, a popular proxy for firm climate finance transition risks, measures the level of greenhouse gas [GHG] emissions associated with a firm's business activities or products. Corporate carbon footprints provide an indication of how much anthropogenic carbon a company contributes to atmospheric GHGs and to global warming (Wiedmann, 2009; Harangozo & Szigeti, 2017). Carbon footprints are preferred by academics and industry practitioners over other climate transition risk metrics, largely because they can be easily converted to dollar losses (using the effective carbon price) or hidden costs (using the future costs of carbon) (BNP Paribas, 2016). Carbon footprints help facilitate the implementation of divestment strategies or low-carbon indices (e.g., S&P Carbon Efficient Indices, MSCI Low Carbon Index) by establishing a link between climate- and financial- risk. Although corporate carbon footprints are a popular metrics in assessing climate transition risks, carbon emissions data has numerous problems, including limited, inconsistent and inaccurate reporting (Kennett et al., 2021; Nguyen et al., 2021).

The GHG Protocol (WRI and WBCSD, 2020) divides carbon emissions into three categories: *Scope 1* - direct emissions from sources and assets controlled by the firm, *Scope 2* - indirect emissions from purchased electricity, and *Scope 3* - indirect emissions from a firms' value chain. Traditionally, the protocol requires all firms to report Scope 1 and Scope 2 emissions, whereas firms have the discretion on *whether* and *which categories* they choose to report Scope 3 emissions. Recently, there are some signals that a mandatory disclosure of Scope 3 emissions is on the way. For instance, a draft rule by the U.S. Securities and Exchange Commission in March 2022 proposes that firms need to disclose emissions generated by their suppliers or partners if they are material or included in any of their emission targets.[3] Therefore, the importance of accurately quantifying Scope 3 emissions is indisputable. Even so, systematically accounting for all emissions along the entire value chain (sometimes up to ten thousand firms) to the same level of accuracy is broadly acknowledged to be extremely challenging (Patchell, 2018).

Scope 3 has many merits - it covers all of the indirect emissions spanning a firm's full value chain, from acquiring and pre-processing raw materials (upstream) to distributing, storing, using, and disposing of the end products sold to customers (downstream). It captures a significant proportion of many firms' total carbon footprints, especially for those operating in the energy sector (Cumberlege, 2013; Hertwich & Wood, 2018). Further, Scope 3 represents the most significant emission reduction opportunities going forward, and a full assessment of Scope 3 emissions is critical to understanding the end-to-end impacts of carbon tax and climate policies on individual firms (Downie & Stubbs, 2013). However, the analyses of firm-level emissions by external stakeholders are usually limited to Scope 1

[3] See https://www.reuters.com/legal/litigation/us-sec-set-unveil-landmark-climate-change-disclosure-rule-2022-03-21/ for more information.

and Scope 2 emissions (Boermans & Galema, 2017; Goldhammer *et al.*, 2017; Griffin *et al.*, 2017). This is due to the following three issues associated with Scope 3:

[1] *No Regulation and Lack of Clear Guidance*. Despite the recent signals, there are no binding rules on Scope 3 emissions disclosure. Further, related sustainability reporting standards/frameworks such as Global Reporting Initiative, Sustainability Accounting Standards Board, and International Integrated Reporting either remain silent on Scope 3 emissions reporting or fail to provide detailed recommendations on how Scope 3 should be properly disclosed (Klaaßen & Stoll, 2021). As measurement and disclosure of Scope 3 are inconsistent and unsystematic, the quality and accuracy of firms' *voluntary* disclosures remain unclear. Given the complexity in calculating Scope 3 (especially when there is no proper guidance apart from the GHG protocol) and extensive data collection efforts are involved, it is not surprising that few firms are willing to disclose Scope 3 emissions.

[2] *Incomplete Composition/ Activity Exclusion*. Firms are not required to disclose the *full composition* of Scope 3 emissions. Thus, using the aggregated Scope 3 emissions data can be misleading. As firms may choose to only report areas that they are performing well or are easier to measure whilst intentionally ignore other areas out of the fifteen distinctive Scope 3 categories.[4] For example, firms that choose to report *Purchased Goods and Services* could have very different value chain emissions and firm characteristics compared to those that only choose to report emissions associated with *Business Travel*. It does not make sense to aggregate emissions data with many missing values, when firms have the discretion to choose which categories they would like to report and the boundaries they would like to report within. Rather than comparing apples with oranges, one should either look at firms' Scope 3 data at the category level, or replace missing values (i.e., unreported Scope 3 categories) with estimated values before performing any cross-sectional comparisons.

[3] *Measurement Divergence/ Reporting Inconsistency*. Firms may set different operational boundaries on the same Scope 3 emissions category, report different values across different communication channels (i.e., annual filings, sustainability reports, or through third-party initiatives such as the Carbon Disclosure Project [CDP]), and/or occasionally update their reported emissions for the past years in later years (Busch & Hoffmann, 2011; Klaaßen & Stoll, 2021). The aforementioned issues make it hard for third-party data providers, such as Bloomberg and Refinitiv, to build consensus and provide consistent Scope 3 measures. To be more specific, third-party data providers may collect Scope 3 emissions data from different sources, update restated values under different time frames, and/or make adjustments to the reported values using different proprietary models. Further, differences

[4] The fifteen categories include: (1) purchased goods and services; (2) capital goods; (3) fuel- and energy-related activities; (4) upstream transportation and distribution; (5) waste generated in operations; (6) business travel; (7) employee commuting; (8) upstream leased assets; (9) downstream transportation and distribution; (10) processing of sold products; (11) use of sold products; (12) end-of-life treatment of sold products; (13) downstream leased assets; (14) franchises; (15) investments. (1)-(8) are generally considered as upstream emissions, whereas (9)-(15) are often classified as downstream emissions. See Section 2 for more details.

in scenarios (e.g., from methodological choices on allocation methods, product use assumptions, end-of-life assumptions) and estimation models make Scope 3 data unreliable and difficult to compare between different data providers (WRI and WBCSD, 2020; Shrimali, 2021). Researcher and industry practitioners (e.g., asset managers and institutional investors) should be aware of the measurement divergence among third-party data providers when performing analysis/forming investment portfolios using Scope 3 data.

In the face of the issues mentioned above for disclosing firms, there is an additional need to develop an estimation model that employs externally available predictors to cover non-disclosing firms in a broader investment universe. Ideally, these models should be implemented top-down using business and financial metrics, as opposed to using a bottom-up process that requires rigorous hand-collected data. Although models using simple extrapolation techniques (CDP, 2016a; MSCI ESG Research, 2016; Thomson Reuters, 2017) or sophisticated machine learning techniques (Han *et al.*, 2021; Nguyen *et al.*, 2021) are readily available for estimating Scope 1 and Scope 2 emissions, little has been done on Scope 3.

A few attempts are discernible from emissions data providers using a variety of modelling approaches. Some providers employ a bottom-up process-based life cycle assessment and use parameters such as firm activities and emission factors (Carbon4Finance).[5] Others employ environmental input-output assessment using top-down metrics at the industry level (Trucost, 2019)[6] or multi-variable regression model using metrics at the firm level (CDP, 2020). Unfortunately, many organisations provide limited information on their Scope 3 estimation methods. Furthermore, there is virtually no disclosure on the prediction performance of these models. The quality/integrity of the estimated Scope 3 datasets is unknown, as evidenced by Busch et al. (2022), who discovered that the correlation between Scope 3 estimations from ISS and Trucost is surprisingly low (16%). The estimation of Scope 3 emissions is important since it helps fill in the gaps (i.e., unreported Scope 3 categories). However, it is problematic that third-party provider estimates do not disclose limitations such as inherent prediction errors and data uncertainties.

Accordingly, using data retrieved from Bloomberg, Refinitiv, and ISS, we examined the following research questions: (i) *What is the quality of Scope 3 emissions data in terms of measurement divergence between data vendors?,* (ii) *What is the quality of Scope 3 emissions data in terms of the composition of emission categories reported by firms?*; and (iii) *Can machine learning methods be used to improve prediction accuracy of Scope 3 emissions for non-disclosing firms?* To answer the first question, we looked at the Scope 3 emissions datasets from Bloomberg/Refinitiv/ISS, and the divergence among

[5] See Anquetin et al (2022) for a detailed description of the methodology of Carbon4Finance.

[6] Trucost only provides upstream Scope 3 data in its data products and only collects disclosure of Scope 3 emissions from "Transport and distribution" since this category is the most common among all upstream categories. For all other categories, they employ the environmental input-output models to estimate emissions data.

these data providers via a three-way reconciliation on aggregated Scope 3 emissions. To answer the second question, we analysed the composition of Scope 3 emissions from Bloomberg (which is the only dataset that has detailed breakdown by categories) and explored the relevance and completeness of each Scope 3 emissions category. To address the third question, we evaluated whether the Scope 3 emissions values could be estimated using top-down business and financial data and whether prediction accuracy could be improved using sophisticated machine learning techniques.

Our study makes several contributions to the existing literature. First, we extend the study by Busch et al. (2022) that analyses the *divergence* in third-party carbon emission datasets for Scope 1, Scope 2 and Scope 3 between 2005 and 2016. Focusing solely on Scope 3 emissions data from 2013 to 2019, we extend beyond correlation analysis to quantify the degree of divergence among raw emissions data and to understand the implication of this divergence on emission ranking. Second, we provide a systematic analysis of the *incomplete composition/ activity exclusion* of Scope 3 emissions reporting by firms for all reporting firms between 2010 and 2019 in the Bloomberg dataset and examine how this problem persists across time and sectors. This is an extension of parts of the Klaaßen and Stoll (2021) analysis, who only examine 56 technology firms in 2019. Finally, to the best of our knowledge, we are the first to apply machine-learning algorithms to predict the aggregated Scope 3 emissions and its individual categories from the set of top-down business and financial data. The prediction accuracy from this framework is indicative of the trustworthiness level of estimated Scope 3 emissions data.

Our main results are summarized as follows. First, we find that there is considerable *divergence* of Scope 3 emissions data among third-party data providers. When the data provider adjusts reported emission values with its proprietary models (in this case, ISS), none of its data points are *identical* to Bloomberg or Refinitiv Eikon (within 1% error), and the *correlation* values of this dataset with the two other datasets are low (55% - 56%.). However, when the data providers use purely reported emission values without any adjustments (in this case, Bloomberg and Refinitiv Eikon), they still have a surprisingly low proportion of *identical* data points (only 68%) despite high correlation values (95%). This divergence will have little impact on the formation of low-carbon portfolios if fund managers rank firms using Bloomberg or Refinitiv Scope 3 emission datasets, yet will lead to substantially different portfolio constituents using ISS Scope 3 adjusted emissions values. This divergence makes it difficult for investors to understand their portfolios' real exposure to climate risks.

Second, we find that firms normally disclose *incomplete* composition of Scope 3 emissions (3.75 out of 15 categories), but they are reporting more categories over time. The most relevant Scope 3 emissions categories differs both between and within industries. *Business Travel* has been reported by most firms (up to 84% in our sample) despite making up less than 1% of the total Scope 3 emissions. Other, more material Scope 3 categories, such as *Use of Sold Products* (make up to 66% of the total Scope 3 emissions) and *Processing of Sold Products* (make up to 8% of the total Scope 3 emissions), have been largely ignored (disclosed up to 18% and 6%, respectively). A simple fill-in-the-gap analysis suggests that if firms report the *full composition* of Scope 3, their total Scope 3 emissions figure could be 60% bigger than currently reported.

Third, Scope 3 prediction accuracy is low, even with the most sophisticated machine learning algorithms and an extensive set of business and financial predictors. In general, it is easier to predict upstream emissions than downstream emissions. Critically, estimating total Scope 3 emissions from category level instead of aggregated level improves prediction accuracy (i.e., mean absolute error [MAE] of log-transformed emissions reduced by 25% in Linear Forest). This is most probably because the aggregated Scope 3 emissions are distorted by non-reported categories, suggesting that the modelling of Scope 3 emission should be conducted at category level. However, there are limited improvements in prediction performance of sophisticated machine learning models (i.e. *Linear Forest*) relative to baseline models (i.e. *Industry Fill* or *Ordinal Least Square*). More precisely, Linear Forest is slightly better at predicting total Scope 3 emissions at category level and aggregated level than baseline models (MAE is reduced by 2% to 6%), and yields more or less equivalent prediction accuracy to a Stepwise regression model across most individual categories. Further, predictor importance varies by category materially. Overall, our findings imply that researchers and investors should be wary of the potential prediction errors when using Scope 3 emissions obtained from third parties. The findings also call for more transparent disclosure from third-party data providers in terms of estimation methodologies and prediction performance.

The rest of the paper proceeds as follows: Section 2 provides context on the Scope 3 emissions problem. Section 3 outlines the data used and Section 4 presents the methodology implemented for the analysis. Section 5 reports the results and Section 6 concludes.

2. Context: The Scope 3 problem

2.1. Accounting and reporting of Scope 3 emissions

The accounting and reporting of Scope 3 emissions (or 'value chain' emissions) largely follow the *GHG Protocol Corporate Value Chain Accounting and Reporting Standard* (WRI and WBCSD, 2020). The protocol differentiates Scope 3 emissions into 15 distinct categories of upstream and downstream emissions. These categories are designed to be mutually exclusive to prevent double-counting, yet firms within the same supply chain or across different supply chains may include the same source of emissions in their Scope 3 reporting. Each category includes several activities that may emit GHG emissions individually, for which a minimum *operational boundary* is established to ensure that major sources of emissions are accounted for (e.g., *cradle-to-gate* or *Scope 1 and Scope 2 emissions* of relevant value chain partners). Further, the mismatch in the timing of firm activities may exist between the firm and its value chain partners. For instance, emissions related to purchased goods may occur before the firm's reporting year, employee commuting may occur simultaneously, and use of sold products may occur

Table 1 – Scope 3 emissions categories

Notes: This table summarises the definitions, operational boundaries and time boundaries of the fifteen Scope 3 emission categories as defined by the *GHG Protocol Corporate Value Chain Accounting and Reporting Standard* (WRI and WBCSD, 2020).

Emissions categories	**Definition**	**Operational boundaries**	**Time boundaries**
Upstream categories			
1. Purchased goods and services	Extraction, production and transportation of goods and services acquired by firm	All upstream (*cradle-to-gate*) emissions of purchased goods and services	Past year, reporting year
2. Capital goods	Extraction, production and transportation of capital goods acquired by firm	All upstream (cradle-to-gate) emissions of purchased goods and services	Past year, reporting year
3. Fuel and energy-related activities (not included in Scope 1 or Scope 2)	Extraction, production and transportation of fuels/energy acquired by firms and not accounted in Scope 1 and 2 (upstream emissions of purchased fuels, electricity, transmissions and distributions loss, generation of purchased electricity to end-users for electricity firm/energy retailers)	For upstream emissions of purchased fuels: All upstream (cradle-to-gate) emissions of purchased fuels For upstream emissions of purchased electricity: All upstream (cradle-to-gate) emissions of purchased fuels For T&D losses: All upstream (cradle-to-gate) emissions of energy consumed in a T&D system For a generation of purchased electricity that is sold to end-users: Emissions from the generation of purchased energy	Past year, reporting year
4. Upstream transportation and distribution	Transportation and distribution of products purchased by the firm between its tier 1 suppliers and its own operation, transportation and distribution of services purchased by the firm (inbound logistics, outbound logistics, between company activities)	The scope 1 and scope 2 emissions of transportation and distribution providers	Past year, reporting year
5. Waste generated in operations	Disposal and treatment of waste generated in firms' operations	The scope 1 and scope 2 emissions of waste management suppliers that occur during disposal or treatment	Reporting year, future year
6. Business travel	Transportation of employees for business-related activities during the reporting year	The scope 1 and scope 2 emissions of transportation carriers that occur during the use of vehicles	Reporting year

7. Employee commuting	Transportation of employees between their homes and worksites	The scope 1 and scope 2 emissions of employees and transportation providers that occur during the use of vehicles	Reporting year
8. Upstream leased assets	Operations of assets leased by the firm	The scope 1 and scope 2 emissions of lessors that occur during the reporting company's operation of leased assets (e.g., from energy use)	Reporting year
Downstream categories			
9. Transportation and distribution	Transportation and distribution of products sold by firms between its operations and end consumers (including retails and storage)	The scope 1 and scope 2 emissions of transportation providers, distributors, and retailers that occur during the use of vehicles and facilities	Reporting year, future year
10. Processing of sold products	Processing of intermediate products sold by downstream companies	The scope 1 and scope 2 emissions of downstream companies that occur during processing	Reporting year, future year
11. Use of sold products	The end-use of goods and services sold by the company	The direct use-phase emissions of sold products over their expected lifetime (i.e., the scope 1 and scope 2 emissions of end-users that occur from the use of products that directly consume energy (fuels or electricity) during use; fuels and feedstocks; and GHGs and products that contain or form GHGs that are emitted during use)	Reporting year, future year
12. End-of-life treatment of sold products	Waste disposal and treatment of products sold by the company at the end of their life	The scope 1 and scope 2 emissions of waste management companies that occur during disposal or treatment of sold products	Reporting year, future year
13. Downstream leased assets	Operations of assets owned by firms and leased to other firms	The scope 1 and scope 2 emissions of lessees that occur during operation of leased assets (e.g., from energy use).	Reporting year
14. Franchises	Operations of franchises	The scope 1 and scope 2 emissions of franchisees that occur during the operation of franchises (e.g., from energy use)	Reporting year
15. Investment	Operations of investment (equity, debt and project finance)	See the description of category 15 (Investments) in section 5.5 for the required and optional boundaries	Reporting year, future year

long after. Therefore, firms should also set a *time boundary* when calculating Scope 3 emissions (See Table 1).

Firms have the discretion to choose which emissions categories to report and whether they would like to extend beyond the minimum boundary to include optional activities. Firms' choices are generally based on five principles: *relevance*, *completeness*, *consistency*, *transparency*, and *accuracy*.[7] There are potential trade-offs among these principles. For instance, firms may choose to report a certain category not because it is material (*relevance*) but because emissions data is easier to collect compared to other categories (*completeness*). To determine which emissions categories and data types to report, firms should identify activities that are most relevant to their businesses and are associated with most GHG emissions. Firms are expected to justify their rationales behind reporting certain emissions categories whilst ignoring others.

Firms generally need two kinds of information to quantify Scope 3 emissions: (i) *activity data*, which represents the level of activities that leads to GHG emissions (e.g., litres of fuel consumed, kilograms of material purchased); and (ii) *emissions factors* that convert quantified activities to GHG emissions (e.g., CO_2 emitted per litre of fuel consumed or per kilogram of material produced). Activity data can be sourced from primary channels (e.g., data obtained directly from suppliers that relate to specific activities in the reporting firm's value chain) or secondary channels (e.g., industry-average data, financial data, proxy data). Primary data is generally considered to be more accurate and more specific to the activities whose emissions are being calculated. Using primary data, as opposed to secondary data, imposes responsibility on particular firms and allows for differentiation in terms of carbon profile between firms. However, under certain circumstances, secondary data can be used to supplement primary data to achieve *completeness* (Patchell, 2018). When primary data is not available, firms may conduct a simple extrapolation to derive its emissions from industry-averages using spend-based metrics. Reporting firms may also perform cascade calculations on how much each of their value chain partners contribute to their total emissions. The most appropriate emissions factor(s) and method(s) used for calculating emissions vary between categories. Table A1 in the Appendix lists all possible calculation methods using the first Scope 3 emissions category - *Purchased Goods and Services* - as an example.

2.2. Selection bias and data errors in Scope 3 emissions

Prior literature has shown that firms' decisions on whether to report Scope 3 emissions, which categories to report and what operational boundaries to establish are affected by many factors. First, while the disclosure of Scope 1 and Scope 2 has been improving rapidly, the disclosure of Scope 3 emissions remains patchy (Matisoff et al., 2013; Fickling & He, 2020). Bigger emitters, including some

[7] See the GHG Protocol for the full definitions of these principles.

unlisted oil and gas firms, are less likely to report Scope 3 emissions and/or offer downstream value chain partners supplier-specific data. Firms are not always able to source data directly from their suppliers. As such, carbon auditing the entire value chain can be a very daunting and costly task (Patchell, 2018). Consequently, firms are less likely to include Scope 3 in their carbon reduction targets, citing that these emissions occur outside of its control.[8]

Second, firms tend to cherry-pick which Scope 3 categories to disclose. Firms are expected to map out all emissions categories in their value chains and identify which ones to include based on relevance and materiality. Yet, they have the motivation to knowingly understate or neglect certain Scope 3 emissions categories. According to the CDP (2018), only 26.7% of the disclosing firms calculate all emissions categories that they consider to be relevant, and this problem is even more prominent if firms report Scope 3 emissions via channels that are under public scrutiny (e.g., corporate reports) (Depoers *et al.*, 2016; Klaaßen and Stoll, 2021). The inconsistency in reported Scope 3 emissions across different communication channels is also known as *reporting inconsistency* (Klaaßen & Stoll, 2021). There are two other sources of errors in Scope 3 emissions, namely *boundary incompleteness* and *activity exclusion* (Klaaßen & Stoll, 2021). *Boundary incompleteness* often arises when firms are not able to source primary or secondary data in a systematic way across various value chain partners or third-party data providers. *Activity exclusion* arises when firms intentionally exclude relevant emissions categories/ business activities in their Scope 3 emissions estimates. An example of *activity exclusion* would be that most reporting firms choose not to disclose emissions from *Purchased Goods and Services* and *Use of Sold Products*, though these are generally considered as the most material emissions categories for firms across different industries (CDP, 2016b). These three sources of errors, if not dealt with carefully, would not only lead to inaccurate emissions calculations, but also make the comparison across different reporting firms difficult.

These problems are intensified, as Scope 3 emissions data has been collected by third-party data providers from different reporting channels and adjusted using different estimation models. Busch et al. (2022) investigated the consistency of emissions data among third-party data providers (including Bloomberg, CDP, ISS, MSCI, Sustainalytics, Thomson Reuters Refinitiv, and Trucost) spanning the period 2005 to 2016. The authors found that the divergence in reported Scope 3 is much more substantial than that of Scope 1 and Scope 2. For instance, the Pearson correlation between ISS and Trucost is surprisingly low (16%). Further, the inconsistencies among data providers tend to grow over time. Part of the reasons for the divergence in Scope 3 is the variation in estimation approaches employed by third-party providers (e.g., process analysis versus input-output analysis), though these methods are expected to produce similar, if not identical, estimation results.

[8] It is expected that this trend will be reversed in the future. For instance, the Science Based Targets initiative (SBTi) requires that if Scope 3 emissions represent more than 40% of their carbon footprint, then firms should set a target to cover this impact (see https://sciencebasedtargets.org/resources/files/SBT_Value_Chain_Report-1.pdf)

3. Data

As noted in the introduction, this paper explores the quality of Scope 3 emission data in terms of divergence and composition and the performance of machine-learning models used to predict Scope 3 emissions (see Section 4.1, Section 4.2 and Section 4.3 respectively for more details).

To address our first research question and see whether divergence exists among data providers, we obtained firm Scope 3 emissions from three sources; ISS, Refinitiv Eikon, and Bloomberg. We study 2013-2019 as it had the most complete data across the datasets. We used ISIN and reporting year to match data points across all three data providers and ended up with 6,725 firm-year observations of *aggregated* Scope 3 emissions values (see Table 2). Refinitiv Eikon and Bloomberg obtained firm *raw* reported Scope 3 emissions from different channels (e.g., the CDP report, firm's annual filings or sustainability reports). ISS uses proprietary modelling and trust methods to provide a wider universe of emissions than just reported CDP data and other publicly available sources. They distinguish different data sources for Scope 1 and 2, including 'CDP', 'Sustainability/Annual Reports', 'Other Reported' and 'Modelled'. We limited our analysis to reported data only since ISS overwrite low trust data with modelled emissions. All Scope 3 emissions are modelled due to the inconsistencies in reporting and are differentiated into upstream and downstream emissions (ISS Methodology, Factset).

To address our second and third research questions, we used Bloomberg Scope 3 emissions data and its categorical breakdown. These 15 distinct categories are defined by the GHG Protocol (see Section 2.1 and Table 1), as well as a miscellaneous category named 'Other' which captures emissions that are not able to be classified into one of these fifteen pre-defined categories. While Scope 3 emissions data is available since 2005 from Bloomberg, the overall number of disclosures is very small before 2010. Therefor we restricted our sample period to 2010-2019. We started with 12,097 *aggregated* Scope 3 firm-year observations, out of which 9,518 have breakdown details (see Table 3). We used both *aggregated* and the detailed breakdowns of Scope 3 at category level as target variables. For our machine learning prediction analyses, the baseline predictor set contains two financial metrics – total revenues and number of employees – that have been commonly used in the past literature. We also extended the original predictor set to include financial metrics from firm's annual income statements and balance sheets (see Section 4.3 for more details). As *relevant* Scope 3 emissions categories differ significantly between sectors, we divided firms into smaller groups using their GICS group codes so that varied emission patterns across industries could be properly reflected. All financial predictors and industry classifications are retrieved from Refinitiv Eikon and are matched back to Bloomberg emissions dataset using ISIN and reporting year. Our final sample for Scope 3 predictions consists of 11,109 firm-year observations. 988 firm-year observations were removed due to no breakdown details, values being missing, extremely small or large, or non-normally distributed.

Summary statistics of our final dataset, for the parts of the analysis that addresses the third research questions, is presented in Appendix Table A5.

4. Method

4.1. Data Quality: Divergence

We started by examining the *divergence* between three of the largest third-party datasets (ISS ESG, Refinitiv Eikon and Bloomberg). Busch et al. (2022) applied the Pearson/Spearman correlation analyses to measure firm carbon emissions consistency among third-party providers. We went further by seeking to quantify the degree of divergence among emissions data and to understand the implication of data inconsistency on emission ranking (emissions ordered from highest to lowest per provider).

To do this, we obtained aggregated Scope 3 emission data from the three data providers and calculated the proportion of data points that are 'identical' (i.e., with an absolute percentage error of less than 1%, see Equation 1 below) between all three providers, as well as the proportion of data points that are not identical but within an acceptable error range (i.e., with an absolute percentage error of less than 20%, see Equation 2 below).

$$\%Identical_{AB} = \frac{1}{n} * \sum \left\{ 1 if \left| \frac{Emission_{it}^{B}}{Emission_{it}^{A}} - 1 \right| \leq 1\% \right\} \tag{1}$$

$$\%Acceptable_{AB} = \frac{1}{n} * \sum \left\{ 1 if \left| \frac{Emission_{it}^{B}}{Emission_{it}^{A}} - 1 \right| \leq 20\% \right\} \tag{2}$$

where n is the number of overlapped firm-year observations between dataset A and dataset B, $Emission_{it}^{B}$ and $Emission_{it}^{A}$ are the aggregated Scope 3 emissions for firm i in reporting year t obtained from dataset A and dataset B, respectively.

We further investigated whether the divergence in Scope 3 emissions has a substantial effect on emission ranking. This is particularly relevant in the construction of low-carbon portfolios (e.g., S&P Carbon Efficient Indices, MSCI Low Carbon Indices),[9] where rating agencies and/or investors may overweigh firms in lower emission deciles whilst underweight firms in higher emission deciles. To do so, all firm-year observations obtained from Bloomberg, Refinitiv Eikon and ISS were assigned to different ranking deciles based on emissions, and the proportion of observations that stay in the same or adjacent ranking deciles was identified using Equation 3 and Equation 4, respectively:

[9] See https://www.spglobal.com/spdji/en/landing/investment-themes/carbon-efficient/ and https://www.msci.com/msci-low-carbon-indexes

$$\%SameDecile_{AB} = \frac{1}{n} * \sum \left\{ 1 if Decile_{Emission_{it}^{B}} = Decile_{Emission_{it}^{A}} \right\} \tag{3}$$

$$\%AdjDecile_{AB} = \frac{1}{n} * \sum \left\{ 1 if Decile_{Emission_{it}^{B}} = Decile_{Emission_{it}^{A}} \pm 1 \right\} \tag{4}$$

where $Decile_{Emission_{it}^{A}}$ and $Decile_{Emission_{it}^{B}}$ are the respective ranking deciles dataset A and dataset B's Scope 3 emissions fall into, respectively.

4.2. Data Quality: Composition

In the second part of our analysis, we investigate the quality of the *composition* of Bloomberg's Scope 3 emissions. For each category, we measured *relevance* based on its relative contribution to the firm's aggregated Scope 3 emissions profile, and *completeness* based on the proportion of firms that choose to disclose this category.

Carbon intensity was calculated using Equation 5. Normalizing by total revenues allows us to compare Scope 3 emissions across firms of different sizes:

$$Intensity_{it}^{Cat} = \frac{Emission_{it}^{Cat}}{\Re V_{it}} \tag{5}$$

where *Cat* represents one of the fifteen Scope 3 emissions categories. $Intensity_{it}^{Cat}$, $Emission_{it}^{Cat}$, and $\Re V_{it}$ represent carbon intensity, Scope 3 emissions, and total revenues of firm *i* in reporting year *t*, respectively.

The relative contribution of each emissions category to the *full composition* of Scope 3 was measured by Equation 6. Note that we set unreported categories (i.e., missing carbon intensity values) to zero. As a result, the contribution of any unreported category for firm *i* in reporting year *t* would be zero.

$$Contribution_{it}^{Cat} = \frac{Intensity_{it}^{Cat}}{Intensity_{it}^{Scope3}} \tag{6}$$

The *completeness* of each Scope 3 emissions category was then calculated as:

$$\%Disclosure^{Cat} = \frac{1}{n} * \sum_{it} \{ 1 if Disclosure_{it}^{Cat} = 1 \} \tag{7}$$

where $Disclosure_{it}^{Cat}$ is a dummy variable taking the value of one if a specific Scope 3 emissions category is reported by firm *i* in year *t*.

Finally, to calculated 'corrected' Scope 3 emissions, we substituted unreported Scope 3 categories by the median carbon intensity of all other firms in the nearest available peer group (see Equation 8 below):

$$Emission_{it}^{Correct} = Emission_{it} + \sum_{Cat} \left\{ Intensity_{IND,t}^{Cat} * \Re V_{it} if Disclosure_{it}^{Cat} = 0 \right\} \tag{8}$$

where $Emission_{it}$ is the raw and incomplete Scope 3 emissions data obtained from Bloomberg for firm *i* in year *t*, and $Emission_{it}^{Correct}$ is the 'corrected' emissions figure by filling in all missing values using peer group median. $Intensity_{IND,t}^{Cat}$ is the median carbon intensity for all reporting firms in firm *i*'s nearest available peer group. Here, 'nearest available peer group' generally refers to firms within the same GICS sub-industry. We require the peer group to have at least 10 firms. When there are not sufficient observations from the same GICS sub-industry, we gradually extended our criteria to include firms operating within the same GICS industry, GICS industry group, and finally, GICS sector.

4.3. Machine Learning Prediction Methods

In the third part of the analysis, for non-disclosing firms, we develop estimation models to predict aggregated Scope 3 emissions and the fifteen distinct categories that make up Scope 3 emissions. We followed Nguyen et al. (2021), who employed machine learning algorithms to predict firms' Scope 1 and Scope 2 emissions from a set of externally available data.

4.3.1. Baseline Models

We started with two baseline models for benchmarking purposes. The first model is an *Industry Fill* model, in which the *aggregated* Scope 3 emissions and its individual categories were estimated using the median of disclosed emissions data of the firm's nearest available peer group, similar to Section 4.2. For each Scope 3 emissions category, we estimated non-disclosing firm's carbon emissions using Equation 9. All notations carry the same meaning as that of Equation 8.

$$Emission_{i \in IND,t}^{Cat} = Intensity_{IND,t}^{Cat} * \Re V_{it} \tag{9}$$

After each individual category has been estimated, aggregated Scope 3 emissions were calculated as:

$$Emission_{it} = \sum_{Cat} Emission_{i \in IND,t}^{Cat} \tag{10}$$

The second baseline model is a simple Ordinal Least Square [OLS] regression that predicts Scope 3 emissions at categorical level using two financial metrics, namely, Revenue ($\Re V_{it}$) and Total Employees (Emp_{it}), and a set of dummy industry indicators (IND) for j GICS groups. Both emissions values and predictor values were transformed using natural logarithm to account for non-normal distributions.

$$logEmission_{it}^{Cat} = f(log\Re V_{it}, logEmp_{it}, \sum_j IND) \tag{11}$$

4.3.2. Linear Models

There could be other financial metrics that are better at capturing Scope 3 emission patterns across the entire supply chain. In Nguyen et al. (2021), the set of predictors chosen for Scope 1 and Scope 2 emissions includes: Revenue, Total Assets, Number of Employees, Intangible Assets, Net Property Plant and Equipment [NPPE], Capital Expense, Gross Margin, Leverage and Capital Intensity. We therefore extended our original predictor set to include additional financial variables such as: Cost of Goods Sold, Earnings Before Interest and Taxes [EBIT], Earnings Before Interest, Taxes, Depreciation, and Amortization [EBITDA], Operational Expense, Net Income, Total Debt, Current Asset, Current Liability, Inventory, and Receivables.

Given the limited number of observations in the Scope 3 emissions dataset, the employment of the extended set of predictors could lead to multicollinearity issues. To avoid this, we employed the *Forward-Backward Stepwise Regression*, which automatically includes relevant predictors (< 1% significant level) into the model and excludes irrelevant ones (>5% significant level). The list of top five relevant predictors for each Scope 3 emissions category could be found in Table A6 of the Appendix.

We further employed *Elastic Net*, a penalized linear regression model to address potential multicollinearity issues (Zou & Hastie, 2005). The regularization strategies are to shrink the size of the coefficients on identical predictors. This is achieved by adding two penalty terms, λ_1 and λ_2, to the sum of squared estimate of errors [SSE] as weights on the sum of squared coefficients and the absolute values of the coefficients. If the coefficient estimates are inflated by multicollinearity, they would be shrunken down to 1/k of one single predictor (use sum of squared coefficients) or an absolute zero (use the absolute value of coefficients) it they are inflated by multicollinearity. The penalty terms are usually referred to as hyper-parameters in machine learning algorithms and are optimized using Bayesian hyperparameter optimization in the training set. We employ five-fold cross-validation to optimize the hyperparameters as well as to compare the performance across different prediction models (the model is optimized and trained on four folds and is evaluated on the mean error of the remaining hold-out fold). All yearly observations of a firm are either included in the same training subset or in the holdout set, so that prediction performance is evaluated on the *non-disclosing* firms.

4.3.3. Tree Based Ensemble Models

Application of the *Tree-Based Ensembles* has been reported in recent emission modelling literature to yield superior prediction performance compared to other modelling techniques (Han *et al.*, 2021; Nguyen *et al.*, 2021). This is because these models can capture non-linearity and correlations among predictor sets, and at the same time, improve stability and interpretability of coefficients in predicting carbon emissions. We employed two types of tree-based ensemble models in this paper – namely, *Random Forest* and *Extreme Gradient Boosting.*

The basis of tree-based ensemble is decision trees (Breiman et al., 1984). A decision tree includes multiple branches using if-then statements, and the estimation of the target variable is calculated using constant approximation (i.e. the mean value of observations in the same branch). The if-then statement is formed by choosing a predictor and its split value by minimizing the best aggregated SSE from two sub-samples. The tree either grows into maximum depth or is "pruned" to shallow trees. Several trees could be combined into an ensemble, in which predictions are combined individually via *Random Forest* (RF) (Breiman, 2001) or sequentially via *Extreme Gradient Boosting* (XGB) (Chen & Guestrin, 2016). The main hyperparameters for these models are the number of trees and the maximum depth. Potential overfitting problems are addressed by restricting the maximum depth of the trees (so as not to create too complex decision boundaries) and by growing multiple trees with randomness (by subsampling predictors or subsampling observations). For the XGB model, we also optimized other hyper-parameters (e.g., minimum child weight, column sample by trees, subsample and regularized alpha). The hyperparameters are continued to be optimized using Bayesian hyperparameter optimization in five-fold cross validations.

4.3.4. Linear Tree Models

Linear Tree models are a special form of tree-based models with a linear functional model in each leaf. Linear tree perfectly combines the learning ability of decision trees and the extrapolation power of linear models. Thus, the hybrid model leads to better predictive power and better insights than either model alone. Linear Forests generalize Random Forests algorithms by combining linear models with the same Random Forests. An initial linear model is fitted on the whole dataset, then the residuals from this model are used as a target variable for the subsequent Random Forest, and the final predictions are generated by using the sum of predictions from the initial linear model and the residual predictions from the Random Forest (Zhang et al., 2019).

Similar to XGB, Linear Boosting builds models subsequentially in a two-stage process. Starting with an initial linear model, a simple decision tree is fitted to model the residuals of the previous steps. During this process, the branch with the highest absolute predicted residual is identified and a binary vector is identified based on the observations to this rule. This binary vector is then fitted into the initial linear model until a certain stopping criterion is met (Ilic et al., 2021). The hyperparameters remain similar to tree-based ensemble models and are optimized using Bayesian hyperparameter optimization in five-fold cross validations.

5. Result

5.1. Data Quality: Divergence

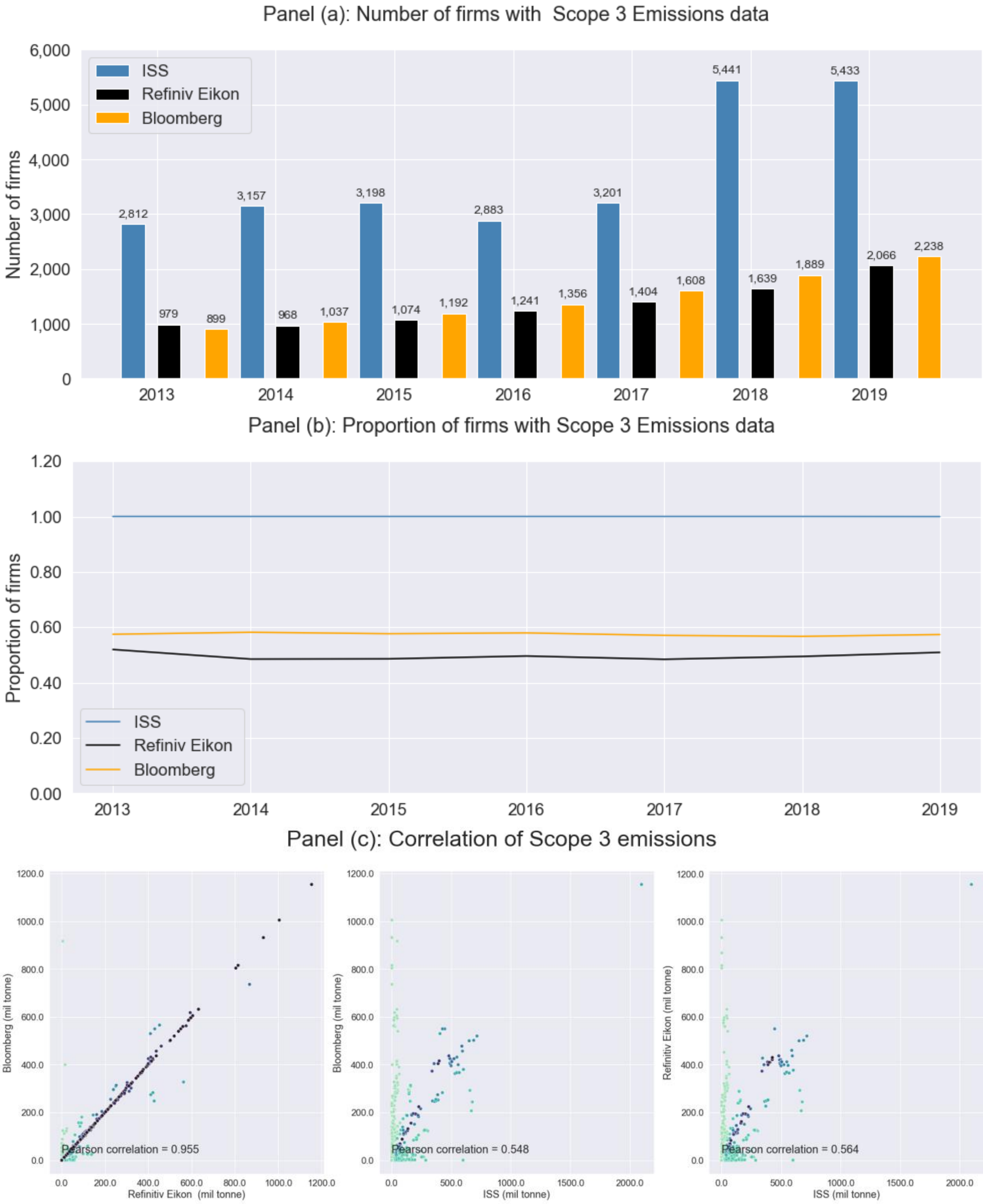
This section presents the result from the divergence analysis detailed in Section 4.1. *Figure 1* presents the coverage of Scope 3 among the three data providers over time. ISS uses proprietary models to fill missing values for unreported firms and has the largest coverage among the three providers (5,433 firms as of 2019), followed by Bloomberg (2,238 firms as of 2019) and Refinitiv Eikon (2,066 firms as of 2019; see Panel a). We observed a steady increase in the sample size of all three data providers over time, except for ISS, who have had a sizeable expansion in 2018. On average, 57% - 60% of firms that report Scope 1 and/or Scope 2 emissions in Bloomberg and Refinitiv Eikon also report Scope 3 emissions, and this proportion remains steady throughout the entire sample period.[10] For ISS, the proportion of firms reporting Scope 3 is 100% of that provide Scope 1/Scope 2 emissions metrics (see Panel b). This is not surprising, given that ISS adjusts all reported/missing Scope 3 emissions data with their own estimates to address the inconsistencies in reporting and to differentiate between upstream and downstream emissions.

Panel (a) of *Table 2* presents the reconciliation results on Scope 3 emissions. We find that the fraction of *identical data* among Bloomberg, Refinitiv Eikon and ISS is surprisingly low. While it is expected that *none* of the ISS adjusted values are within the 1% error range of the same firm-year observation obtained from the other two datasets, the low proportion of *identical* data points (68%) between Bloomberg and Refinitiv was unexpected given that *reported* emissions are supposed to be similar, especially when some are extracted directly from the same communication channel (i.e., corporate reports or CDP). This problem persists even when the cut-off error range is extended to 20%. We find that while 84% of the data points between Refinitiv Eikon and Bloomberg are within an *acceptable range* (20% cut-off), only 5% of ISS's Scope 3 emissions is similar to the other two datasets. Despite this, the Pearson pairwise correlations between the three datasets (Bloomberg and Refinitiv Eikon, Bloomberg and ISS, and Refinitiv Eikon and ISS) are relatively high - reaching an average of 95%, 55%, and 56%, respectively.

[10] Specifically, the ISS data set have 26,127 data points on Total Emissions, 26,065 data points on Scope 1 Emissions, 26,043 data points on Scope 2 emissions and 26,125 data points on Scope 3 emissions. The Refinitiv Eikon dataset have 15,305 data points on total emissions, 13,074 data points on Scope 1 emissions, 12,833 data points on Scope 2 emissions and 8,027 data points on Scope 3 emissions. The Bloomberg dataset have 17,812 data points on Scope 1 emissions, 15,918 data points on Scope 1 emissions, 15,590 data points on Scope 2 emission sand 10,219 data points on Scope 3 emissions.

Figure 1 – Scope 3 data coverage over time

Notes: This figure summaries the data coverage of the three Scope 3 emission datasets (Bloomberg, Refinitiv Eikon and ISS) in 2013-2019. Panel (a) presents the number of observations over time, panel (b) presents the proportion of firms that also disclose Scope 3 emissions beside Scope 1 or 2 emissions, and panel (c) presents the pairwise scatter plots of the three datasets.



[11]

[11] Prior to this analysis, we remove two data outliers for Bloomberg (detected due to plotting) (see Appendix Figure 2 for the details of outliers and Appendix Table 1 for Correlation before this removal)

Table 2 –Divergence between three Scope 3 emission datasets in 2013-2019

Notes: This table summarises the analyses of divergence between three of the largest third-party datasets (ISS ESG, Refinitiv Eikon and Bloomberg) as described in Section 4.1. The three-way matched dataset includes 6,725 firm-year observations of aggregated Scope 3 emissions in 2013-2019. Panel (a) presents the divergence in raw emission values in proportion of identical data, proportion of data within acceptable error range and correlation over time. Panel (b) presents the divergence in emission ranking in proportion of data within the same decile ranking and within the adjacent rankings.

Panel (a): Divergence in 'raw' emissions values

Year	Obs	Proportion of identical data (<1% measurement error)			Proportion of data within acceptable error range (<20% measurement error)			Pearson Correlation		
		Bloomberg vs Refinitiv Eikon	Bloomberg vs ISS	ISS vs Refinitiv Eikon	Bloomberg vs Refinitiv Eikon	Bloomberg vs ISS	ISS vs Refinitiv Eikon	Bloomberg vs Refinitiv Eikon	Bloomberg vs ISS	ISS vs Refinitiv Eikon
2013	614	67%	0%	0%	82%	4%	4%	99%	64%	65%
2014	710	69%	1%	1%	85%	4%	5%	99%	65%	65%
2015	798	68%	0%	0%	84%	5%	5%	86%	56%	63%
2016	882	69%	0%	0%	84%	5%	5%	89%	47%	51%
2017	1046	67%	0%	0%	83%	5%	5%	97%	46%	45%
2018	1249	70%	0%	0%	84%	6%	6%	99%	47%	47%
2019	1426	67%	0%	0%	83%	6%	6%	99%	63%	62%
Total	**6725**	**68%**	**0%**	**0%**	**84%**	**5%**	**5%**	**95%**	**55%**	**56%**

Panel (b): Divergence in emissions ranking

Year	Obs	Proportion of data within the same ranking			Proportion of data within adjacent rankings		
		Eikon vs Bloomberg	ISS vs Bloom-berg	ISS vs Eikon	Eikon vs Bloomberg	ISS vs Bloomberg	ISS vs Eikon
2013	614	80%	22%	21%	95%	54%	56%
2014	710	85%	23%	24%	96%	54%	55%
2015	798	82%	23%	22%	96%	55%	55%
2016	882	82%	23%	23%	96%	55%	56%
2017	1046	81%	21%	22%	95%	54%	54%
2018	1249	82%	22%	22%	95%	53%	52%
2019	1426	83%	23%	23%	94%	50%	50%
Total	**6725**	**82%**	**22%**	**22%**	**95%**	**53%**	**53%**

We find that emission rankings are more consistent among three datasets than absolute emissions values (*Panel b – Table 2*). Specifically, Bloomberg has 82% (13%) of data points that are in the same (adjacent) ranking decile(s) as Refinitiv Eikon. ISS consistently differs in that only 22% (31%) of its emissions data falls into the same (adjacent) ranking decile(s) as the other two datasets. This implies that low-carbon indices/portfolios constructed from ISS would differ significantly from those constructed using emissions data obtained from Bloomberg or Refinitiv.

There seems to be a clear time effect and size effect in determining the level of inconsistency among these data providers (Table A2 in the Appendix). First, we observe that ISS becomes increasingly inconsistent with both Refinitiv Eikon and Bloomberg, while the other two datasets converge over time.[12] Second, we observe that most of the divergence happens in firms with extremely low (bottom 10%) or extremely high (top 90%) emissions values, or firms with extremely low (top 10%) or extremely high (top 90%) total revenues. Further, data inconsistency is more prominent among sectors that are relatively green (e.g., Real Estate, Financials, Information Technology, Communication Services and Consumer Staples) or those that are extremely brown (e.g., Utilities) between Refinitiv Eikon and Bloomberg.[13]

Overall, we show that there is considerable divergence in Scope 3 emissions among third-party data providers, especially when the data provider (in this case, ISS) adjusts values using its proprietary estimation models. This divergence persists over time and is more prominent for firms that are extremely large or extremely small (as measured by total revenues and emissions).

5.2. Data Quality: Composition

This section presents the result from the data composition analysis detailed in Section 4.2. Table 3 summarises the composition of Scope 3 using data obtained from Bloomberg. From 2010 to 2019, 4,590 firms (identical to 23,782 firm-year observations) report at least one emissions ‘scope’, of which, 57% report *aggregated* Scope 3 emissions (2,502 firms or 12,097 firm-year observations) and 42% provides the detailed breakdown of Scope 3 emissions by category (1,972 firms or 9,518 firm-year observations; see Table 3, *Columns 1 and 2*).

From 2010 to 2019, all reporting firms in Bloomberg emit a total of 52.5 gigatons for Scope 1, 9.7 gigatons for Scope 2, and 134.3 gigatons for Scope 3 (Table 3, *Column 3*). We compared these figures to the total energy-related GHG emissions (321.5 gigaton from 2010 to 2019)[14], and discovered that Bloomberg only covers around one-third of the total direct emissions worldwide. On average, firms

[12] A similar analysis is conducted for each emission scopes and is available upon request. Generally, the result is similar for Scope 1 and Scope 2 emissions. For Scope 3 emissions, as ISS data is modelled thus no abnormal trend can be detected.

[13] In terms of geographic locations [Panel (c)], we observe the lowest level of consistencies among European countries. This is observed across three datasets. Breakdown by the scope of emissions, we realize that this divergence is largely driven by inconsistencies in Scope 2 (see Table 6).

[14] https://www.iea.org/articles/greenhouse-gas-emissions-from-energy-data-explorer

Table 3 – Summary of Scope 3 emissions composition from the Bloomberg dataset in 2010-2019

Notes: This table summarises the Scope 3 emission composition from the Bloomberg dataset in 2010-2019, as described in Section 4.2. Panel (a) presents emission values by scopes, panel (b) presents the detailed breakdown of Scope 3 emissions.

Emissions Values	**(1) Number of firm-year obs**	**(2) Number of firms**	**(3) Combined emissions among all disclosing firms (CO_2-e tonne)**	**(4) Average emissions among disclosing firms (CO_2-e tonne)**	**(5) Proportion of disclosure**
Panel (a): Emission Scopes					
Total Emissions	21,166	4,183	72,742,408,565	3,436,757	
Scope 1	18,529	3,854	52,496,207,234	2,833,192	88%
Scope 2	18,136	3,777	9,787,471,240	539,671	86%
Scope 3	12,097	2,502	134,304,221,822	11,102,275	57%
Scope 3 (aggregated of 15 activities)	9,518	1,972	114,373,422,867	12,016,539	45%
Panel (b): Scope 3 Emissions Categories					
Upstream Categories					
1. Purchased Goods and Service	3,033	814	9,670,931,758	3,188,570	32%
2. Capital Goods	1,631	472	813,850,793	498,989	17%
3. Fuel and Energy Related Activities	2,995	782	6,399,788,737	2,136,824	31%
4. Upstream transportation and Distribution	2,462	640	984,597,167	399,918	26%
5. Waste Generation	3,182	818	107,408,065	33,755	33%
6. Business Travel	8,035	1,607	893,687,913	111,224	84%
7. Employee Commuting	3,092	776	186,371,542	60,275	32%
8. Upstream Leased Assets	944	255	20,841,366	22,078	10%
Downstream Categories					
9. Downstream Transportation & Distribution	2,087	500	908,767,627	435,442	22%
10. Process of Sold Products	616	170	6,836,013,095	11,097,424	6%
11. Use of Sold Products	1,680	433	75,799,055,679	45,118,486	18%
12. End of Line Treatment	1,251	349	967,348,056	773,260	13%
13. Downstream Leased Assets	778	217	126,888,409	163,096	8%
14. Franchise	524	138	242,345,701	462,492	6%
15. Investments	730	203	1,136,129,408	1,556,342	8%
16. Other	3,752	750	12,700,645,592	3,385,033	39%

emit 2.8 million tons of Scope 1, 0.5 million tons of Scope 2 and 11.1 million tons of Scope 3 greenhouse gases (Table 3, *Column 4*). A firm's emissions profile is largely captured by Scope 3, even if the emissions values are normalised by total revenues to avoid potential size biases.

Figure 2 and Figure 3 exhibit the *completeness* and *relevance* of Scope 3 emission over time and across different GICS sectors, respectively (The underlying numbers are reported in Table A3 and Table A4 in the Appendix). On average, firms only report 3.75 out of the 15 distinct Scope 3 emissions categories during our sample period. The degree of *completeness* is relatively low. However, see significant improvements over time – the average number of reported Scope 3 categories (4.69) in 2019 has tripled that in 2010 (1.72), and we see a significant increase in the proportion of disclosing firms in most Scope 3 emissions categories. Further, firms tend to report categories that are easier to calculate, but not those that are more material to their organisations' carbon footprint. Firms report emissions related to business travels (84%) more than any other categories during our sample period, despite the fact that *Business Travel* only covers less than 1% of the total emissions along the value chain. While most Scope 3 emissions could be captured by *Use of Sold Products* (75.7 gigatons or 63%), less than 20% report this emissions category (1,680 firm-year observations or 18%). The second and the third largest emissions categories - *Purchased Goods and Services* and *Capital Goods* – are also largely ignored.

The most *relevant* Scope 3 categories vary greatly across GICS sectors. For most firms (especially those operating in *Energy* and *Consumer Discretionary* sectors, which includes oil & gas firms and fuel-based car manufacturers), a significant portion of Scope 3 emissions comes from *Use of Sold Products*. In sharp contrast however, firms operating in the *Financials* sector fund emissions via their loan/ investment portfolios. Consequently, most of their Scope 3 emissions come from *Investments*. For firms operating in the *Health Care* and *Consumer Staples* GICS sectors, most of their Scope 3 emissions come from *Purchased Goods and Services*, whereas for *Utilities* firms, *Fuel and Energy Related Activities* contributes the most to Scope 3 emissions.

Finally, we calculate the 'corrected' Scope 3 emissions by substituting unreported Scope 3 categories by the median carbon intensity of all other firms in the nearest available peer group following Equation 8. This fill -in-the-gap analyses are conducted on the subset of firms that disclose incomplete emission composition with available revenue data. A simple fill-in-the-gap analysis suggests that if firms report the *full composition* of Scope 3, their total Scope 3 emissions figure could be 60% higher than currently reported. Detailed analysis by year and sector could be found in Figure A1 and A2 in the Appendix.

Figure 2 – Completeness and Relevance of Scope 3 emission categories over time

Notes: This figure summaries the analyses of completeness and relevance of Scope 3 emission categories over time as described in Section 4.2. Panel (a) presents the completeness based on the proportion of firms that choose to disclose each category, panel (b) presents the relevance based on the relative contribution of each category to Scope 3 intensities. Unreported categories (i.e., missing carbon intensity values) are set to zero. The sample is 9,518 observations from 1,972 firms that disclose the composition of Scope 3 in the Bloomberg dataset in 2010-2019.

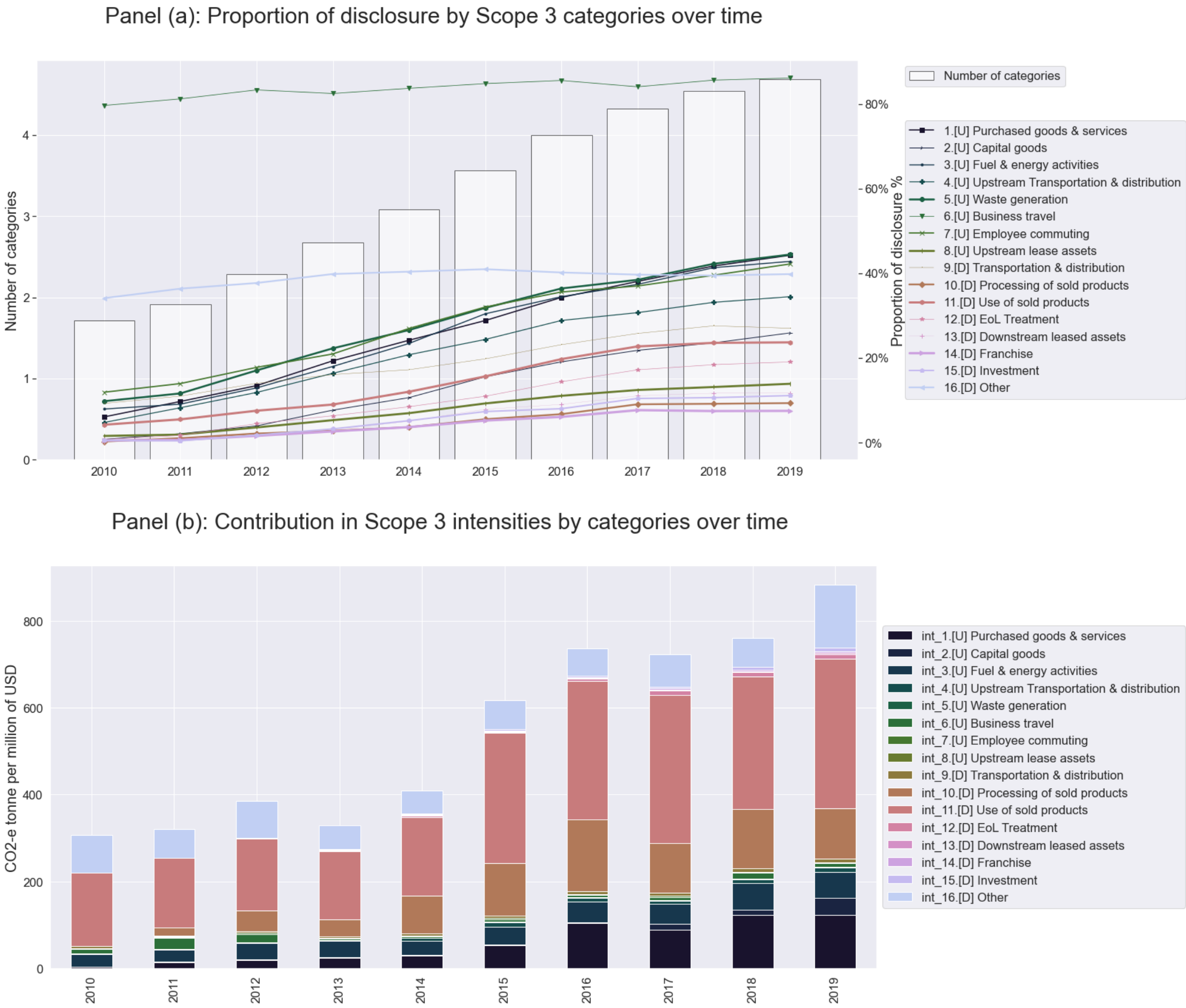
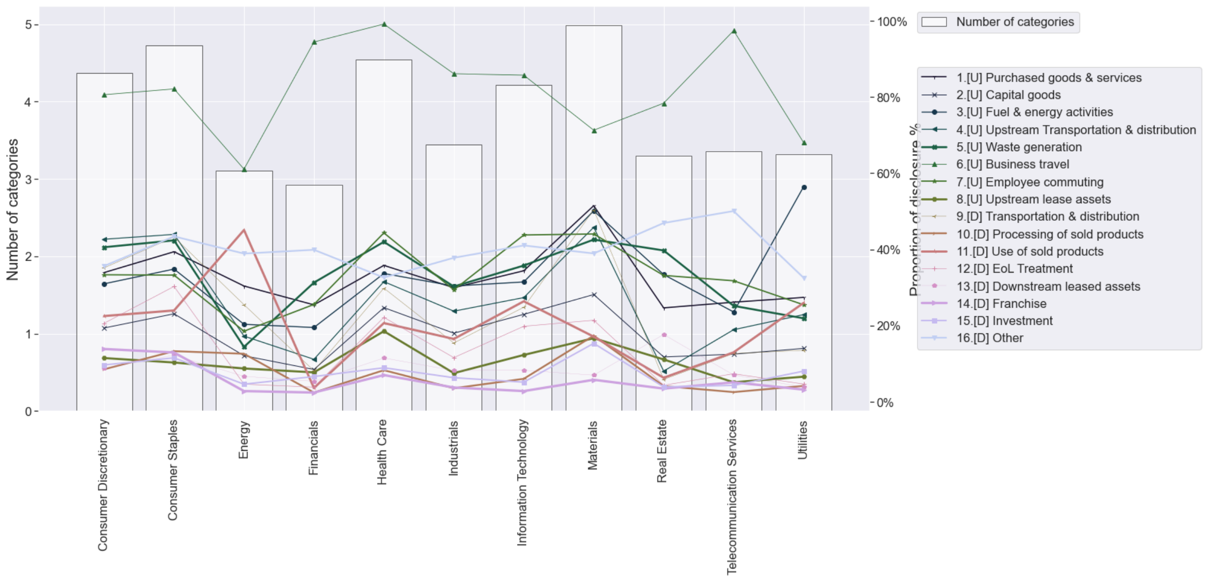


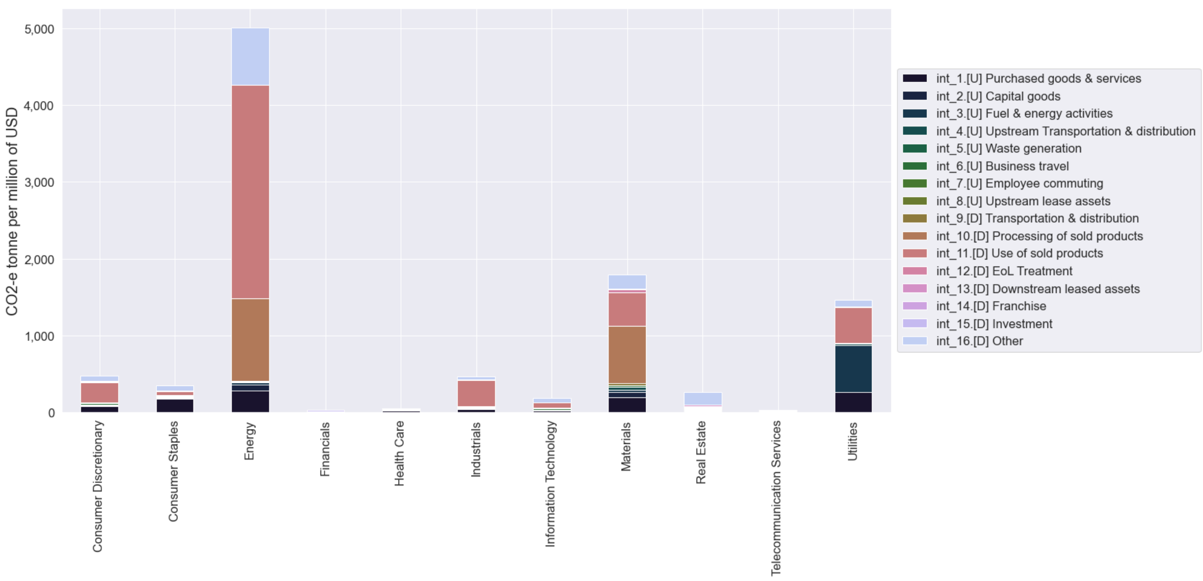
Figure 3 – Completeness and Relevance of Scope 3 emissions category by sectors

Notes: This figure summaries the analyses of completeness and relevance of Scope 3 emission categories by GICS sectors as described in Section 4.2. Panel (a) presents the completeness based on the proportion of firms that choose to disclose each category, panel (b) presents the relevance based on the relative contribution of each category to Scope 3 intensities. Unreported categories (i.e., missing carbon intensity values) are set to zero. The sample is 9,518 observations from firms that disclose the composition of Scope 3 in the Bloomberg dataset in 2010-2019.

Panel (a): Proportion of disclosure by Scope 3 categories by sectors



Panel (b): Contribution in Scope 3 intensities by categories by sectors



5.3. Performance of Machine Learning Prediction Models

Table 4 presents the out-of-sample prediction performance of all models presented in Section 4.3. The main criterion of performance assessment is the mean absolute error (MAE) of log-transformed emissions in five-fold cross-validation (the model is optimized and trained on four folds and is evaluated on the mean error of the remaining fold). All yearly observations of a firm are either included in the same training subset or in the holdout set, so that prediction performance is evaluated on the *non-disclosing* firms. In panel (a), we compare prediction results on aggregated Scope 3 emissions when they are treated as a single value (that is, only one machine learning model is built) and they are aggregated from a group of 16 sub-models made up from 15 categories (see Table 1) and the residual covered by the 'Other' category. In panel (b), we present prediction results on individual categories.

First, we find that our two baseline models (the *industry-fill* model (Table 4, Column 1) and the *naïve OLS model* (Table 4, Column 2)) produce very similar prediction performance on both the *aggregated* Scope 3 emissions and its individual categories. However, the *industry-fill* model underperforms *naïve OLS* in emissions estimates for certain categories (e.g., *Franchise*, *Investment*, *Downstream Transportation and Distribution*, and *Downstream Leased Asset*) when we only have limited reported emissions data.

Second, we find that the application of machine learning algorithms to be more useful when each category is estimated individually and aggregated into the total Scope 3 emissions values. Both the *industry fill* model and the *naïve OLS* model generate a log-MAE of 1.88 when *aggregated* values are used to predict total Scope 3 emissions. Given that firms generally report incomplete compositions, we first obtained emissions estimates from all sixteen categories, and then aggregated them into total Scope 3 emissions. By doing so, we see a large improvement in prediction accuracy for both models, as evidenced by a decline in log-MAE from 1.88 to 1.35 for the *industry fill* model and from 1.88 to 1.34 for *naïve OLS*.

Third, we find that it is easier to predict *upstream emissions* than *downstream emissions*. The best prediction performance is found in *Business Travel* (*naïve OLS* log-MAE: 0.99), *Employee Commuting* (*naïve OLS* log-MAE: 1.18) and *Capital Goods* (*naïve OLS* log-MAE: 1.27). There are two possible reasons behind this: (i) firms report more of the emissions associated with these categories, and (ii) good proxies could be found in suppliers' financial statements that help capture emissions derived from these upstream activities. For instance, number of employees could be used to calculate emissions associated with business travel/employee commuting, whereas capital expenditures (CAPEX) might be a good indicator for emissions associated with capital goods.

We extended beyond the baseline models by including (i) all possible financial predictors that may capture Scope 3 emissions alongside the supply chain (*full OLS* - Table 4, Column 3), and (ii) the most relevant (i.e., significant) predictors chosen by forward-backward stepwise (*stepwise OLS* - Table

Table 4 – Performance of machine learning prediction models

Notes: This table summarises the out-of-sample performance of machine learning prediction models for Scope 3 emission as described in Section 4.3. The primary criterion is the mean absolute error (MAE) of log-transformed emissions averaged from five-fold divisions. Panel (a) presents the prediction results on aggregated Scope 3 emissions when they are treated as a single value and when they are aggregated from a group of 16 sub-models (for 15 categories and other). Panel (b) presents the prediction results on individual categories.

Emission Values	Number of Observations			Models								
	Training	Testing	All	(1) Industry Fill	(2) Naïve OLS	(3) Full OLS	(4) Stepwise OLS	(5) Elastic Net	(6) XGBoost	(7) Random Forest	(8) Linear Boost	(9) Linear Forest
Panel (a): Emission Scopes												
Scope 3	8,887	2,222	11,109	1.88	1.88	1.86	1.85	1.90	1.80	1.83	1.86	1.77
Scope 3 (aggregated of 15 activities)	7,043	1,761	8,804	1.35	1.34	1.40	1.31	1.34	1.37	1.35	1.35	1.32
Panel (b): Scope 3 Emission Categories												
Upstream Categories												
1. Purchased Goods and Service	2,178	545	2,723	1.71	1.72	1.73	1.69	1.73	1.79	1.78	1.77	1.74
2. Capital Goods	1,100	275	1,375	1.42	1.27	1.24	1.15	1.21	1.29	1.24	1.32	1.25
3. Fuel and Energy Related Activities	2,161	540	2,701	1.44	1.48	1.46	1.52	1.45	1.43	1.51	1.47	1.41
4. Upstream transportation and Distribution	1,726	432	2,158	1.56	1.48	1.52	1.47	1.52	1.49	1.52	1.57	1.48
5. Waste Generation	2,173	543	2,716	1.44	1.45	1.47	1.42	1.45	1.48	1.52	1.51	1.40
6. Business Travel	5,647	1,412	7,059	0.96	0.99	0.94	0.94	0.97	0.96	0.95	0.96	0.94
7. Employee Commuting	2,205	551	2,756	1.22	1.18	1.17	1.11	1.16	1.13	1.15	1.12	1.11
8. Upstream Leased Assets	358	90	448	2.27	1.78	1.98	1.66	1.78	1.78	2.04	1.99	1.75
Downstream Categories												
9. Downstream Transportation & Distribution	1,430	357	1,787	1.55	1.50	1.51	1.43	1.47	1.62	1.55	1.66	1.47
10. Process of Sold Products	238	60	298	3.10	2.44	2.47	2.38	2.54	2.47	2.82	2.76	2.55
11. Use of Sold Products	1,117	279	1,396	1.63	1.63	1.72	1.61	1.62	1.88	1.85	1.69	1.66
12. End of Line Treatment	744	186	930	1.81	1.71	2.05	1.61	1.73	1.87	1.98	1.74	1.73
13. Downstream Leased Assets	264	66	330	2.65	2.08	2.46	1.91	1.94	2.14	2.25	2.18	2.04
14. Franchise	102	26	128	2.47	1.95	3.59	1.60	2.12	2.01	2.18	1.95	2.20
15. Investments	294	73	367	3.12	2.30	2.76	1.87	2.18	2.22	2.46	2.34	2.27
16. Other	2,597	649	3,246	2.57	2.35	2.40	2.29	2.34	2.41	2.38	2.31	2.32

4, Column 4). While *stepwise* significantly improves prediction accuracy, whether *full OLS* outperforms *naïve OLS* remains inconclusive.

Surprisingly, machine learning algorithms only lead to limited improvements in prediction accuracy compared to the baseline methods (industry fill and naïve OLS models) as well as the best OLS model (Stepwise model). Out of all machine learning techniques used (*Elastic Net*, *XGBoost*, *Random Forest*, *Linear Boost*, *Linear Forest*), only *Linear Forest* (see Table 4, Column 9) consistently outperforms baseline models in predicting total Scope 3 emissions (using both aggregated data and category-level data) (MAE is reduced by 6% and 2%, respectively). *Linear Forest* is slightly better at predicting the aggregated Scope 3 emissions than the Stepwise model (MAE is reduced from 1.85 to 1.77 – 4%) but yields more or less equivalent prediction accuracy across most individual emissions categories (MAE is 1.32 as compared to 1.31 from Stepwise regression). However, the gain in prediction performance by using state-of-the-art machine learning algorithms is very limited when there is poor data quality and low observations in certain categories.

When doing stepwise regressions, we find that predictor importance varies by category materially. For certain emissions categories (e.g., *Employee Commuting*, *Use of Sold Products*, and *Upstream Leased Assets*), *total revenues* is the most important size factor. For other categories, *total revenues* may not be relevant, and other financial metrics might be better proxies for calculating emissions. For instance, both *Purchased Goods and Services* and *Use of Sold Products* are better estimated if *level of inventory* is included in the estimation model, whilst *Capital Goods* is better captured by *capital expenditures* figure in the same reporting year. We also included GICS industry group dummies in the forward-backward stepwise regression. Similar to financial metrics, the importance of industry group indicators varies greatly between categories. For instance, *Insurance*, *Retailing*, *Transportation* and *Materials* are important GICS industry group indicators for *Upstream Leased Assets*, whereas the *Software & Services* is important in predicting emissions associated with *Business Travel*. The top five most relevant predictors for each emissions category could be found in Table A6 in the Appendix.

So far, we measure prediction accuracy on log-transformed data. Thus, it would be interesting to see how this prediction model performs on absolute emissions values (Table 5). To do so, we transformed log-MAE (Table 5, Column 2) into median absolute percentage error [MDAPE] on non-transformed emissions values (Table 5, Column 3) and calculated the proportion of emissions estimates that lies within +/- 50% of the actual emissions values [PPAR] (Table 5, Column 4).

We find that despite the gain in prediction accuracy (especially when each category is estimated individually and aggregated into total Scope 3 emissions values), even the best prediction model has substantial prediction error. Specifically, the stepwise regression model generates a median absolute percentage error of 72.3% on *aggregated* Scope 3 emissions values (by combining all categorical-level

Table 5 – Performance of the best model with alternative measures

Notes: This table summarises the out-of-sample performance of the best prediction model for Scope 3 emission (Stepwise OLS) as described in Section 5.3. The main criterion is the mean absolute error (MAE) of log-transformed emissions averaged from five-fold divisions. Median Absolute Percentage Error (MDAPE) and proportion of prediction in acceptance range (PPAR) are used as the alternative performance criteria on raw emission values. The calculations of alternative measures are presented below. Here $\hat{y}$ is the predicted value of the log-transformed target variable y, $\bar{y}$ is the mean of log-transformed actual y, n is the number of observations in the dataset, e is the exponential, i is the firm and t is the reporting year.

$$MAPE = median\frac{\left|e^{Y_{i,t}} - e^{\widehat{Y_{i,t}}}\right|}{e^{Y_{i,t}}}, \qquad PPAR = \frac{1}{n} * \sum\left(1\ if\ \frac{\left|e^{Y_{i,t}} - e^{\widehat{Y_{i,t}}}\right|}{e^{Y_{i,t}}} \leq 50\%, 0\ if\ otherwise\right) * 100\%$$

Emission Values	MAE (log-scaled)	MDAPE (%)	PPAR (+/-50%)
Panel (a): Emission Scopes			
Scope 3	1.85	92.35	18.60
Scope 3 (aggregated of 15 activities)	1.31	72.23	30.36
Panel (b): Scope 3 Emission Categories			
<u>Upstream Categories</u>			
1. Purchased Goods and Service	1.69	81.59	22.86
2. Capital Goods	1.15	60.27	33.63
3. Fuel and Energy Related Activities	1.52	80.26	24.79
4. Upstream transportation and Distribution	1.47	80.89	25.69
5. Waste Generation	1.42	83.37	26.56
6. Business Travel	0.94	59.65	38.18
7. Employee Commuting	1.11	66.60	32.38
8. Upstream Leased Assets	1.66	83.64	20.63
<u>Downstream Categories</u>			
9. Downstream Transportation & Distribution	1.43	79.99	22.56
10. Process of Sold Products	2.38	92.89	16.01
11. Use of Sold Products	1.61	83.93	25.90
12. End of Line Treatment	1.61	84.46	21.67
13. Downstream Leased Assets	1.91	96.08	21.24
14. Franchise	1.60	187.74	25.41
15. Investments	1.87	89.97	24.24
16. Other	2.29	95.75	18.06

emissions estimates). This means that only 30% of the predicted emissions values lies within +/- 50% of the actual reported emissions values. Potential biases in the reported datasets (ie the training set) may contribute to this outcome to a certain extent. Researchers and industry practitioners should be wary of prediction errors when doing risk analysis using machine learning techniques and/or ‘adjusted’ Scope 3 emissions obtained from third-party data providers.

6. Conclusion

This paper explores the *divergence* and the *composition* of Scope 3 emissions data among third-party data providers (Bloomberg, Refinitiv Eikon, ISS). We find that Scope 3 emissions vary greatly across third-party data providers, with the divergence between ISS (adjusted values) and the other two datasets (reported values) being most prominent. Surprisingly, we find that Bloomberg and Refinitiv only have 68% identical data points, though both datasets rely on firm *reported* data with no further adjustments. While emission rankings are more consistent among three data providers than that of absolute emissions values, only 22% of the ISS’s emissions data falls into the same ranking decile as the other two datasets. Naïve use of ISS data can have a large impact on the formalization of divestment strategies or low-carbon indices.

With respect to the *composition* of Scope 3 emissions, firms tend to only disclose certain emission categories that are easier to calculate, despite these categories only capture a relatively low proportion of total Scope 3 emissions along the value chain. By looking at the relative contribution of each emissions category to the *full composition* of Scope 3, we argue that the most relevant Scope 3 emissions categories vary greatly between industries. Therefore, firms should perform relevant tests to make sure that they establish the ‘correct’ operational boundaries.

The paper also compared the prediction accuracy of various baseline (industry fill, naïve OLS), linear (full OLS, forward-backward stepwise) and machine learning models (elastic net, XGBoost, random forest, linear boost, and linear forest). Overall, prediction is more effective when each emissions category is estimated individually and aggregated into total Scope 3 emissions. Further, we find that predictor importance varies greatly by emissions category, and upstream emissions are easier to predict compared to downstream emissions. Contrary to our expectations, even when we use the most advanced machine learning techniques, the improvements in prediction accuracy are minimal. However, the extent to which potential biases in the reported datasets (ie the training set) contributed to this outcome is unknown.

Overall, our findings emphasize the need for improvement in Scope 3 emissions disclosure. First, binding mandates should be established, and more guidance is needed to derive accurate calculations. Second, firms should expand their operational reporting boundaries to include categories that are the most relevant to their businesses. Third, it is also important for firms to source primary data from value chain partners, as secondary data, especially emissions estimates are subject to large uncertainties. Researchers and industry practitioners should be wary and incorporate the potential errors in publicly disclosed emissions data and the limitations of various prediction models.

References

BNP Paribas. (2016). *Stress-testing equity portfolios for climate change factors: The carbon factor*. Available at http://securities.bnpparibas.com. [Accessed 10 Mar. 2018]

Boermans, M., & Galema, R. (2017). Pension funds' carbon footprint and investment trade-offs. *DeNederlandscheBank Working Paper*. No. 554.

Breiman, L. (2001). Random forests. *Machine learning*. Vol. 45, No. 1, pp. 5-32.

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and regression trees* (Vol. 37).

Busch, T., & Hoffmann, V. H. (2011). How hot is your bottom line? Linking carbon and financial performance. *Business & Society*. Vol. 50, No. 2, pp. 233-265.

Busch, T., Johnson, M., & Pioch, T. (2022). Corporate carbon performance data: Quo vadis? *Journal of Industrial Ecology*. Vol. 26, No. 1, pp. 350-363.

CDP. (2016a). *CDP Clean and Complete Dataset 2016: Technical Annex III: Statistical Framework*. Available at https://b8f65cb373b1b7b15feb-c70d8ead6ced550b4d987d7c03fcdd1d.ssl.cf3.rackcdn.com/comfy/cms/files/files/000/000/633/original/Technical-Annex-III_Statistical-Framework.pdf. [Accessed 04 Jul. 2018]

CDP. (2016b). *Out of the starting blocks Tracking progress on corporate climate action*. Available at https://cdn.cdp.net/cdp-production/cms/reports/documents/000/001/228/original/CDP_Climate_Change_Report_2016.pdf?1485276095. [Accessed 04 Jul. 2018]

CDP. (2018). *How can companies address their scope 3 greenhouse gas emissions?* Available at https://www.cdp.net/en/articles/companies/how-can-companies-address-their-scope-3-greenhouse-gas-emissions{CDP, 2018 #7163}. [Accessed 04 Jul. 2018]

CDP. (2020). *CDP Full GHG Emissions Dataset Technical Annex IV: Scope 3 Overview and Modelling*. Available at https://cdn.cdp.net/cdp-production/comfy/cms/files/files/000/003/076/original/2020_01_23_Scope_3_Overview.pdf. [Accessed 04 Jul. 2018]

Chen, T., & Guestrin, C. (2016). *Xgboost: A scalable tree boosting system.* Paper presented at the Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining.

Cumberlege, T. (2013, 16 Apr. 2018). Make business sense of Scope 3. *CarbonTrust*.Available at https://www.carbontrust.com/news/2013/04/make-business-sense-of-scope-3-carbon-emissions/. [Accessed 20 Apr. 2018]

Downie, J., & Stubbs, W. (2013). Evaluation of Australian companies’ scope 3 greenhouse gas emissions assessments. *Journal of Cleaner Production*. Vol. 56, pp. 156-163.

Fickling, D., & He, E. (2020). The Biggest Polluters Are Hiding in Plain Sight. *Bloomberg*.Available at https://www.bloomberg.com/graphics/2020-opinion-climate-global-biggest-polluters-scope-3-emissions-disclosures/. [Accessed 20 March 2022]

Goldhammer, B., Busse, C., & Busch, T. (2017). Estimating Corporate Carbon Footprints with Externally Available Data. *Journal of Industrial Ecology*. Vol. 21, No. 5, pp. 1165-1179.

Griffin, P. A., Lont, D. H., & Sun, E. Y. (2017). The relevance to investors of Greenhouse Gas Emission disclosures. *Contemporary Accounting Research*. Vol. 34, No. 2, pp. 1265-1297.

Han, Y., Gopal, A., Ouyang, L., & Key, A. (2021). Estimation of Corporate Greenhouse Gas Emissions via Machine Learning. *arXiv preprint arXiv:2109.04318*.

Harangozo, G., & Szigeti, C. (2017). Corporate carbon footprint analysis in practice–with a special focus on validity and reliability issues. *Journal of Cleaner Production*. Vol. 167, pp. 1177-1183.

Hertwich, E. G., & Wood, R. (2018). The growing importance of scope 3 greenhouse gas emissions from industry. *Environmental Research Letters*. Vol. 13, No. 10, p. 104013.

Ilic, I., Görgülü, B., Cevik, M., & Baydoğan, M. G. (2021). Explainable boosted linear regression for time series forecasting. *PATTERN RECOGNITION*. Vol. 120, p. 108144.

Kennett, H., Diaz-Rainey, I., Biswas, P. K., & Kuruppuarachchi, D. (2021). Climate Transition Risk in New Zealand Equities. *JOURNAL OF SUSTAINABLE FINANCE AND INVESTMENT*. Vol. Forthcoming.

Klaaßen, L., & Stoll, C. (2021). Harmonizing corporate carbon footprints. *NATURE COMMUNICATIONS*. Vol. 12, No. 1, pp. 1-13.

Matisoff, D. C., Noonan, D. S., & O'Brien, J. J. (2013). Convergence in environmental reporting: assessing the Carbon Disclosure Project. *BUSINESS STRATEGY AND THE ENVIRONMENT*. Vol. 22, No. 5, pp. 285-305.

MSCI ESG Research. (2016). *Filling the blanks: Comparing carbon estimates again disclosures*. Available at https://www.msci.com/documents/10199/139b2ab7-c95f-4f09-9d33-fdc491c5316e. [Accessed 10 May 2018]

Nguyen, Q., Diaz-Rainey, I., & Kuruppuarachchi, D. (2021). Predicting corporate carbon footprints for climate finance risk analyses: A machine learning approach. *ENERGY ECONOMICS*. Vol. 95, p. 105129.

Patchell, J. (2018). Can the implications of the GHG Protocol's scope 3 standard be realized? *JOURNAL OF CLEANER PRODUCTION*. Vol. 185, pp. 941-958.

Shrimali, G. (2021). Scope 3 Emissions: Measurement and Management. *STANFORD SUSTAINABLE FINANCE INITIATIVE*.

Thomson Reuters. (2017). *ESG Carbon Data and Estimate Models*. Available at https://financial.thomsonreuters.com. [Accessed 16 Apr. 2018]

Wiedmann, T. (2009). Editorial: Carbon Footprint and Input-Output - An introduction. *ECONOMIC SYSTEMS RESEARCH*. Vol. 21, No. 3, pp. 175-186.

WRI and WBCSD. (2020). *Corporate Value Chain (Scope 3) Standard*. World Resources Institute and World Business Council for Sustainable Development Available at https://ghgprotocol.org/sites/default/files/standards/Corporate-Value-Chain-Accounting-Reporing-Standard_041613_2.pdf. [Accessed 10 Mar. 2018]

Zhang, H., Nettleton, D., & Zhu, Z. (2019). Regression-enhanced random forests. *ARXIV PREPRINT ARXIV:1904.10416*.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *JOURNAL OF THE ROYAL STATISTICAL SOCIETY: SERIES B (STATISTICAL METHODOLOGY)*. Vol. 67, No. 2, pp. 301-320.

Appendix

Table A1 - Example of the calculation method for Scope 3 category 1 - Purchased goods and services

Notes: This table summarises the calculation method for Category 1 – Purchased goods and services as defined by the *GHG Protocol Corporate Value Chain Accounting and Reporting Standard* (WRI and WBCSD, 2020).

Activity	**Description**
Supplier-specific method	Collects product-level cradle-to-gate GHG inventory data from goods or services suppliers.
Hybrid method	Uses a combination of supplier-specific activity data (where available) and secondary data to fill the gaps. o Collecting allocated scope 1 and scope 2 emission data directly from suppliers o Calculating upstream emissions of goods and services from suppliers' activity data on the n of materials, fuel, electricity, used, distance transported, and waste generated from the production of goods and services and applying appropriate emission factors o Using secondary data to calculate upstream emissions wherever supplier-specific data is not available.
Average-data method	Estimates emissions for goods and services by collecting data on the mass (e.g., kilograms or pounds), or other relevant units of goods or services purchased and multiplying by the relevant secondary (e.g., industry average) emission factors (e.g., average emissions per unit of good or service).
Spend-based method	Estimates emissions for goods and services by collecting data on the economic value of goods and services purchased and multiplying it by relevant secondary (e.g., industry average) emission factors (e.g., average emissions per monetary value of goods).

Table A2 – Biases in Divergence of Scope 3 Emission Datasets

Notes: This table breakdowns the divergence in raw emission values between three of the largest third-party datasets (ISS ESG, Refinitiv Eikon and Bloomberg) as described in Section 4.1 into Sectors (panel (a)), Emission Deciles (panel (b)) and Revenue Deciles (panel (c)).

Panel (a): Divergence in 'raw' emissions values by sectors

Sector	Obs	Proportion of identical data (<1% measurement error)			Proportion of data within acceptable error range (<20% measurement error)			Pearson Correlation		
		Bloomberg vs Eikon	Bloomberg vs ISS	ISS vs Eikon	Bloomberg vs Eikon	Bloomberg vs ISS	ISS vs Eikon	Bloomberg vs Eikon	Bloomberg vs ISS	ISS vs Eikon
Communication Services	432	71%	0%	0%	87%	4%	5%	94%	29%	29%
Consumer Discretionary	668	71%	0%	0%	84%	4%	3%	99%	66%	66%
Consumer Staples	464	66%	0%	0%	83%	4%	3%	96%	39%	39%
Energy	294	78%	0%	1%	88%	18%	16%	99%	90%	90%
Financials	1349	64%	0%	0%	79%	4%	5%	92%	2%	2%
Health Care	326	64%	0%	0%	85%	8%	7%	92%	58%	58%
Industrials	1160	73%	0%	0%	86%	3%	4%	97%	-1%	-1%
Information Technology	676	69%	1%	1%	84%	7%	6%	99%	25%	25%
Materials	572	70%	0%	0%	86%	8%	9%	90%	13%	13%
Real Estate	369	62%	0%	0%	78%	3%	4%	49%	0%	0%
Utilities	415	64%	0%	0%	83%	4%	5%	96%	38%	38%
Total	**6725**	**68%**	**0%**	**0%**	**84%**	**5%**	**5%**	**95%**	**55%**	**55%**

Panel (b): Divergence in 'raw' emissions values by emission deciles

Scope 3 Emissions Deciles	Obs	Proportion of identical data (<1% measurement error)			Proportion of data within acceptable error range (<20% measurement error)			Pearson Correlation		
		Bloomberg vs Eikon	Bloomberg vs ISS	ISS vs Eikon	Bloomberg vs Eikon	Bloomberg vs ISS	ISS vs Eikon	Bloomberg vs Eikon	Bloomberg vs ISS	ISS vs Eikon
below 10%	554	70%	1%	1%	84%	10%	11%	3%	3%	3%
20%	580	64%	0%	0%	81%	6%	6%	56%	0%	0%
30%	601	72%	0%	0%	85%	4%	4%	96%	5%	5%
40%	717	69%	0%	0%	84%	4%	5%	96%	6%	6%
50%	695	66%	0%	0%	81%	3%	3%	95%	4%	4%
60%	698	63%	1%	1%	80%	4%	4%	84%	3%	3%
70%	687	66%	0%	0%	83%	4%	4%	97%	4%	4%
80%	797	72%	0%	0%	86%	4%	3%	99%	-2%	-2%
90%	703	70%	0%	0%	86%	6%	6%	89%	7%	7%
above 90%	693	71%	0%	0%	85%	9%	9%	94%	51%	51%
Total	**6725**	**68%**	**0%**	**0%**	**84%**	**5%**	**5%**	**95%**	**55%**	**55%**

Panel (c): Divergence in ‘raw’ emissions values by revenue deciles

Revenue Deciles	**Obs**	**Proportion of identical data (<1% measurement error)**			**Proportion of data within acceptable error range (<20% measurement error)**			**Pearson Correlation**		
		Bloomberg vs Eikon	Bloomberg vs ISS	ISS vs Eikon	Bloomberg vs Eikon	Bloomberg vs ISS	ISS vs Eikon	Bloomberg vs Eikon	Bloomberg vs ISS	ISS vs Eikon
below 10%	488	71%	0%	0%	86%	5%	5%	84%	9%	9%
20%	508	75%	0%	0%	88%	7%	7%	100%	17%	17%
30%	571	72%	0%	0%	84%	3%	4%	100%	45%	45%
40%	610	71%	0%	0%	85%	4%	4%	100%	24%	24%
50%	662	67%	0%	0%	83%	4%	4%	96%	13%	13%
60%	744	70%	1%	0%	84%	6%	6%	98%	27%	27%
70%	744	67%	1%	0%	80%	6%	7%	96%	18%	18%
80%	719	65%	0%	0%	83%	6%	6%	99%	14%	14%
90%	846	65%	0%	0%	82%	4%	3%	99%	14%	14%
above 90%	831	65%	0%	1%	83%	7%	8%	93%	63%	63%
Total	**6725**	**68%**	**0%**	**0%**	**84%**	**5%**	**5%**	**95%**	**55%**	**55%**

Table A3 - Completeness and Relevance of Scope 3 emissions over time

Notes: This table summaries the analyses of completeness and relevance of Scope 3 emission categories over time as described in Figure 2.

Panel (a): Completeness (Proportion of disclosure)

FY	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**All**
Average number of categories	**1.72**	**1.92**	**2.29**	**2.67**	**3.09**	**3.56**	**3.99**	**4.33**	**4.54**	**4.69**	**3.75**
Proportion of disclosure %											
Upstream Categories											
Purchased Goods and Services	6.1	9.7	13.5	19.3	24.2	28.8	34.3	38.1	41.7	44.2	31.9
Capital Goods	0.8	2.1	3.9	7.7	10.6	15.5	19.1	21.8	23.6	25.9	17.1
Fuel and Energy Related Activities	7.9	9.1	12.9	18.0	23.4	30.4	34.5	37.4	41.3	42.8	31.5
Upstream transportation and Distribution	4.8	8.2	11.8	16.4	20.8	24.4	28.8	30.7	33.1	34.5	25.9
Waste Generation	9.8	11.6	17.1	22.3	26.6	31.8	36.4	38.4	42.2	44.4	33.4
Business Travel	79.6	81.2	83.3	82.4	83.7	84.8	85.5	84.0	85.6	86.1	84.4
Employee Commuting	11.9	14.0	17.8	20.9	27.0	32.0	35.6	37.0	39.5	42.2	32.5
Upstream Leased Assets	1.6	1.9	3.6	5.3	7.0	9.3	11.1	12.4	13.1	13.9	9.9
Downstream Categories											
Downstream Transportation and Distribution	9.3	11.0	14.0	16.1	17.2	19.8	23.1	25.8	27.6	27.0	21.9
Process of Sold Products	0.3	1.1	2.2	2.8	3.7	5.5	6.7	9.1	9.2	9.3	6.5
Use of Sold Products	4.2	5.5	7.5	9.0	12.0	15.6	19.7	22.7	23.6	23.7	17.7
End of Line Treatment	0.8	1.7	4.5	6.3	8.5	11.0	14.4	17.2	18.4	19.1	13.1
Downstream Leased Assets	0.3	0.6	1.8	2.9	5.1	7.9	9.0	11.1	11.7	11.7	8.2
Franchise	0.5	0.6	1.6	2.7	3.7	5.2	6.1	7.7	7.5	7.5	5.5
Investments	0.5	0.4	1.8	3.2	5.2	7.3	8.0	10.4	10.6	11.1	7.7
Other	34.1	36.4	37.7	39.8	40.4	41.0	40.2	39.7	39.5	39.8	39.4

Panel (b): Relevance (Contribution in Scope 3 intensities (CO_2-e tonne per million dollars in revenue)

FY	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	All
Upstream Categories											
Purchased Goods and Services	2.64	14.09	19.17	23.76	29.31	51.76	103.33	87.84	123.18	123.08	77.87
Capital Goods	0.07	0.12	0.60	0.96	1.45	2.59	2.95	13.96	11.64	38.93	11.93
Fuel and Energy Related Activities	29.34	28.54	38.06	37.41	32.75	40.09	46.92	45.56	60.63	59.28	46.57
Upstream transportation and Distribution	0.74	0.94	1.77	2.00	5.79	10.67	9.18	7.91	9.10	11.00	7.53
Waste Generation	0.31	0.37	0.54	0.59	0.85	1.41	1.54	1.78	1.47	1.58	1.27
Business Travel	11.49	25.16	18.16	2.99	2.83	5.96	4.37	7.18	13.49	7.43	8.78
Employee Commuting	0.51	2.92	2.80	2.33	2.29	2.61	2.18	1.89	1.11	1.57	1.92
Upstream Leased Assets	0.06	0.05	0.18	0.15	0.13	0.13	0.15	0.40	0.22	0.54	0.26
Downstream Categories											
Downstream Transportation and Distribution	5.64	2.42	3.61	3.46	4.04	4.90	5.86	7.29	9.86	9.67	6.68
Process of Sold Products	0.01	18.51	48.47	37.85	87.78	122.61	165.38	114.76	135.56	115.06	103.86
Use of Sold Products	168.37	161.09	164.88	158.44	181.40	299.40	320.75	341.14	305.24	345.25	277.49
End of Line Treatment	0.04	0.17	1.41	1.92	2.33	2.63	4.24	9.89	9.98	9.55	5.91
Downstream Leased Assets	0.01	0.01	0.01	1.64	1.96	1.84	1.64	1.87	3.26	3.91	2.17
Franchise	0.00	0.04	0.72	0.89	1.15	1.03	2.33	3.08	2.37	2.77	1.89
Investments	0.00	0.00	0.08	0.12	2.13	2.66	2.85	3.63	7.51	9.79	4.32
Other	87.32	66.08	84.09	54.65	53.90	67.84	63.31	75.59	66.72	144.12	82.03

Table A4- Completeness and Relevance of Scope 3 emissions by sectors

Notes: This table summaries the analyses of completeness and relevance of Scope 3 emission categories by sectors as described in Figure 2.

Panel (a): Completeness (Proportion of disclosure)

Sector	Consumer Discretionary	Consumer Staples	Energy	Financials	Health Care	Industrials	Information Technology	Materials	Real Estate	Telecommunication Services	Utilities
Average number of categories	**4.37**	**4.73**	**3.11**	**2.92**	**4.55**	**3.45**	**4.21**	**4.99**	**3.30**	**3.36**	**3.31**
Proportion of disclosure %											
Upstream Categories											
Purchased Goods and Services	33.9	39.4	30.4	25.5	35.8	30.0	34.4	51.5	24.7	26.2	27.4
Capital Goods	19.4	23.1	12.1	8.5	24.7	18.0	23.0	28.2	11.8	12.5	14.0
Fuel and Energy Related Activities	31.0	34.9	20.4	19.5	33.7	30.4	31.5	50.1	33.4	23.6	56.5
Upstream transportation and Distribution	42.7	44.0	17.3	11.2	31.6	23.8	27.4	45.8	8.1	18.9	22.9
Waste Generation	40.6	42.4	14.4	31.3	42.0	30.3	35.8	42.6	39.7	25.2	21.9
Business Travel	80.6	82.1	61.1	94.4	99.1	86.1	85.7	71.3	78.4	97.4	68.1
Employee Commuting	33.4	33.3	18.6	25.6	44.3	29.4	43.8	44.1	33.2	31.8	25.5
Upstream Leased Assets	11.5	10.3	8.8	7.7	18.6	7.5	12.3	16.7	11.1	5.1	6.6
Downstream Categories											
Downstream Transportation and Distribution	35.2	43.3	25.5	8.2	29.9	15.5	24.8	50.4	5.9	12.7	13.5
Process of Sold Products	8.6	13.3	12.6	2.5	8.3	3.6	6.1	17.4	4.1	2.6	4.2
Use of Sold Products	22.5	24.0	45.1	3.7	20.7	16.5	26.4	17.1	6.3	13.0	26.0
End of Line Treatment	20.6	30.3	4.6	3.9	22.2	11.6	19.8	21.4	4.4	7.4	4.7
Downstream Leased Assets	8.8	12.3	6.7	5.4	11.5	8.2	8.3	7.0	17.7	7.1	3.9
Franchise	13.9	13.0	2.8	2.5	7.0	3.8	2.8	5.8	3.5	5.1	3.2
Investments	9.7	11.6	4.6	6.7	9.0	6.3	5.2	15.4	3.9	4.3	8.1
Other	35.6	43.4	38.9	39.9	32.6	37.8	41.1	38.9	46.9	50.1	32.6

Panel (b): Relevance (Contribution in Scope 3 intensities (CO_2-e tonne per million dollars in revenue)

Sector	Consumer Discretionary	Consumer Staples	Energy	Financials	Health Care	Industrials	Information Technology	Materials	Real Estate	Telecommunication Services	Utilities
Upstream Categories											
Purchased Goods and Services	85.07	172.35	285.19	2.26	23.88	43.77	26.11	197.73	13.67	8.37	262.89
Capital Goods	3.60	3.71	76.81	1.00	2.75	5.03	3.39	70.46	10.81	2.04	5.11
Fuel and Energy Related Activities	2.74	9.76	28.40	0.37	1.36	8.88	3.44	26.32	10.57	1.50	604.15
Upstream transportation and Distribution	6.37	12.13	2.97	0.12	3.60	6.95	2.32	38.81	0.10	1.04	17.24
Waste Generation	1.85	1.60	1.18	0.04	0.90	1.24	0.19	2.95	5.33	0.12	1.17
Business Travel	18.33	1.53	4.53	7.91	3.61	5.10	20.02	10.92	10.62	2.12	0.74
Employee Commuting	1.22	0.84	0.94	4.70	2.03	1.55	1.81	1.12	1.19	0.61	0.32
Upstream Leased Assets	0.11	0.17	0.46	0.11	0.24	0.10	0.43	0.17	1.52	0.15	0.34
Downstream Categories											
Downstream Transportation and Distribution	4.68	12.77	7.91	0.09	2.43	2.97	1.32	36.44	6.90	0.32	15.22
Process of Sold Products	0.18	6.81	1,075.06	-	0.17	0.73	0.63	736.97	14.11	0.00	0.36
Use of Sold Products	269.78	50.51	2,780.24	0.00	7.82	340.23	71.74	437.46	6.07	6.81	459.39
End of Line Treatment	5.24	7.34	1.58	0.00	0.59	4.54	0.72	42.00	0.03	0.05	8.67
Downstream Leased Assets	0.14	1.61	0.02	0.66	0.05	5.38	0.08	0.45	18.74	0.75	0.18
Franchise	14.28	4.03	0.01	-	0.01	0.26	0.00	0.00	-	0.04	0.06
Investments	0.15	1.29	0.31	14.70	0.21	0.79	0.74	8.84	0.04	0.49	5.42
Other	66.54	61.43	746.34	2.62	2.94	37.61	51.48	180.95	161.96	4.10	82.03

Table A5 - Summary statistics for prediction models

Notes: This table summarises the dataset employed for machine learning prediction models as described in Section 3. Panel (a) is the dataset before pre-processing, panel (b) is the data set after pre-processing to fill in missing values, remove extreme values or log-transformed for non-normally distributed variables.

Panel (a): Data before processing

Variables	No Obs	Min	Max	Mean	Median	Std Deviation
1- Bloomberg Carbon Emissions						
Scope 3	12,097	0.00	2,048.00	11.10	0.05	61.10
1. Purchased Goods and Service	2,980	0.00	165.56	3.18	0.31	9.66
2. Capital Goods	1,608	0.00	73.38	0.50	0.07	2.37
3. Fuel and Energy Related Activities	2,916	0.00	188.70	2.15	0.03	11.09
4. Upstream transportation and Distribution	2,417	0.00	27.80	0.40	0.04	1.61
5. Waste Generation	3,106	0.00	2.27	0.03	0.00	0.14
6. Business Travel	7,652	0.00	81.70	0.10	0.01	1.61
7. Employee Commuting	3,041	0.00	4.22	0.06	0.01	0.25
8. Upstream Leased Assets	915	0.00	1.90	0.02	0.00	0.10
9. Downstream Transportation & Distribution	2,037	0.00	34.63	0.44	0.04	2.02
10. Process of Sold Products	604	0.00	556.46	11.30	0.01	50.57
11. Use of Sold Products	1,648	0.00	1,000.00	44.86	2.07	118.12
12. End of Line Treatment	1,234	-0.56	30.63	0.77	0.03	3.05
13. Downstream Leased Assets	765	0.00	9.60	0.17	0.00	0.87
14. Franchise	512	0.00	8.64	0.45	0.00	1.38
15. Investments	715	0.00	139.00	1.58	0.01	9.38
16. Other	3,569	0.00	259.96	3.28	0.01	19.74
2- Predictors						
Revenue	11,549	-9,085.57	379,136.00	14,594.14	4,711.91	27,984.46
Total Assets	11,550	8.84	2,888,550.00	78,783.54	9,569.40	262,700.44
Number Of Employees	10,628	0.00	798.00	38.17	14.00	65.19
Net Property, Plant and Equipment	11,386	-2,567.45	259,651.00	5,803.69	1,139.05	15,332.83
Capital Expense	11,192	-54.23	41,461.86	1,060.58	239.44	2,826.00
Gross Profit	9,049	-2,446.03	114,986.00	5,092.69	1,516.86	10,231.91
Operational Expense	10,407	-13,489.08	368,652.00	12,859.07	3,937.98	26,161.06
EBIT	11,522	-8,391.00	71,230.00	2,229.45	594.52	4,935.04
EBITDA	11,520	-8,100.20	81,801.00	3,061.43	902.74	6,361.88
Net Income	11,548	-19,194.59	96,749.11	1,191.39	332.40	3,237.80
Intangible Gross*	12,097	0.00	182,596.00	2,399.92	181.05	8,145.12
Total Debt	11,549	0.00	753,752.00	15,126.66	2,228.87	52,413.79
Cost of Goods Sold	11,549	-94.30	334,858.00	7,104.70	1,386.00	19,173.22
Current Asset	9,511	0.48	1,068,830.00	7,644.66	2,216.33	27,514.24
Current Liability	9,510	0.65	1,067,580.00	6,197.31	1,592.19	26,036.76
Inventory	8,454	0.00	58,753.27	1,465.05	397.25	3,156.57
Receivable	9,942	0.00	126,460.00	2,671.45	732.56	6,998.42
Gross Margin	11,528	-0.38	1.69	0.54	0.50	0.30
Leverage	9,174	0.00	2.98	0.25	0.23	0.17
Capital Intensity	11,364	-1.00	11,671.28	1.58	0.23	109.49
Asset Age	9,888	0.00	53,881.67	22.09	12.16	548.70

Panel (b): Data after processing

Variables	No Obs	Min	Max	Mean	Median	Std Deviation
1- Bloomberg Carbon Emissions						
(log) Scope 3	11,109	-3.22	21.44	11.35	10.88	3.48
(log) 1. Purchased Goods and Service	2,730	0.69	18.92	11.95	12.87	3.65
(log) 2. Capital Goods	1,396	0.83	18.11	11.13	11.47	2.38
(log) 3. Fuel and Energy Related Activities	2,701	-1.61	19.06	10.60	10.55	3.22
(log) 4. Upstream transportation and Distribution	2,160	1.10	17.14	10.57	10.78	2.69
(log) 5. Waste Generation	2,722	-0.92	14.28	7.77	7.93	2.75
(log) 6. Business Travel	7,058	0.00	18.22	8.72	8.84	2.05
(log) 7. Employee Commuting	2,766	-1.61	15.26	9.00	9.18	2.21
(log) 8. Upstream Leased Assets	566	-1.61	14.46	8.38	8.88	2.35
(log) 9. Downstream Transportation & Distribution	1,797	0.00	17.36	10.49	10.77	2.62
(log) 10. Process of Sold Products	365	4.54	20.14	12.28	12.04	3.53
(log) 11. Use of Sold Products	1,427	2.30	20.72	14.44	14.99	3.53
(log) 12. End of Line Treatment	978	0.69	17.24	10.50	10.82	3.18
(log) 13. Downstream Leased Assets	470	3.21	16.08	9.10	8.87	2.56
(log) 14. Franchise	211	3.69	15.97	11.02	10.93	2.92
(log) 15. Investments	444	-0.51	18.75	11.08	11.23	3.02
(log) 16. Other	3,258	-2.30	19.38	9.77	9.73	3.76
2- Predictors						
(log) Revenue	11,109	8.94	26.66	22.27	22.28	1.59
(log) Total Assets	11,109	15.99	28.69	23.11	22.97	1.86
(log) Number of Employees	11,109	0.00	13.59	8.71	9.41	3.00
(log) Net Property, Plant and Equipment	11,109	0.00	26.28	20.59	20.86	2.86
(log) Capital Expense	11,109	7.15	24.45	19.18	19.30	2.01
(log) Gross Profit	11,109	0.00	25.47	16.88	20.68	8.66
(log) Operational Expense	11,109	0.00	26.63	19.77	21.91	6.97
(log) EBIT	11,109	0.00	24.99	19.60	20.20	3.96
(log) EBITDA	11,109	0.00	25.13	20.31	20.62	3.03
(log) Net Income	11,109	0.00	25.30	18.04	19.62	5.84
(log) Intangible Gross*	11,109	0.00	25.93	14.43	19.28	9.35
(log) Total Debt	11,109	0.00	27.35	20.76	21.51	4.30
(log) Cost of Goods Sold	11,109	0.00	26.54	18.01	21.15	8.11
(log) Current Asset	11,109	0.00	27.70	18.05	21.17	8.05
(log) Current Liability	11,109	0.00	27.70	17.79	20.82	7.95
(log) Inventory	11,109	0.00	24.80	14.75	19.07	8.70
(log) Receivable	11,109	0.00	25.56	17.78	20.16	6.99
Gross Margin	11,109	0.05	1.00	0.53	0.48	0.30
Leverage	11,109	0.00	0.69	0.23	0.21	0.15
Capital Intensity	11,109	0.00	6.10	0.53	0.23	0.93
Asset Age	11,109	0.49	50.99	14.07	12.13	9.41

Table A6 – The top five most relevant predictors in stepwise regression

Notes: This table summarises the top five most relevant predictors for each emission scope in *Forward-Backward Stepwise Regression*, which automatically includes relevant predictors (< 1% significant level) into the model and excludes irrelevant ones (>5% significant level) as described in Section 4.3.2. Rank of each predictor is determined by the order of which the predictors are included into the predictor set.

Predictors	Total Scope 3	[1] Pur-chased Goods and Service	[2] Capital Goods	[3] Fuel and Energy Related Activi-ties	[4] Up-stream Trans-porta-tion and Distri-bution	[5] Waste Gener-ation	[6] Busi-ness Travel	[7] Em-ployee Commuting	[8] Up-stream Leased Assets	[9] Down-stream Trans-porta-tion and Distri-bution	[10] Process of Sold Prod-ucts	[11] Use of Sold Prod-ucts	[12] End of Line Treat-ment	[13] Down-stream Leased Assets	[14] Franchise	[15] Invest-ments	[16] Other
Revenue	1	2		4	1	4	2	1	1	4		1	1				4
Number of Employees			3				1	2									
Gross Profit							5	4									
Operational Expense								5							2		
Net Property, Plant and Equipment*										1	1					1	
Capital Expense	3		1	1		1											1
Leverage														1			
Asset Age			4														
Total Debt																2	
Receivable			2													5	

Inventory		1		3	2							2	3	3			2
Gross Intangible Asset							3	3									
Current Asset		3															
GMAR												5					
Current Liability										5							
[GICS] Banks																4	
[GICS] Commercial & Professional Services																	
[GICS] Consumer Services															1		
[GICS] Diversified Financials																	3
[GICS] Energy			5								3	3		4			5
[GICS] Food, Beverage & Tobacco		5			5					2							
[GICS] Household & Personal Products													4				
[GICS] Insurance									2								
[GICS] Materials		4		5	3	5			5	3	2		2		3	3	
[GICS] Pharmaceuticals, Biotechnology & Life Sciences															5		
[GICS] Retailing									3								
[GICS] Software & Services						3	4						5				
[GICS] Telecommunication Services																	
[GICS] Transportation									4					5			
[GICS] Utilities				2								4					

Figure A1 - Emission profiles after filling in the gaps for unreported categories over time

Notes: This figure summaries the emission profiles before and after filling in the gaps on all unreported categories for the sample of 9,518 observations from 1,972 firms that disclose the composition of Scope 3 in the Bloomberg dataset in 2010-2019 as described in Section 4.2.

Panel (a): Contribution of Scope 3 categories before filling in the gaps

Panel (b): Contribution of Scope 3 categories after filling in the gaps

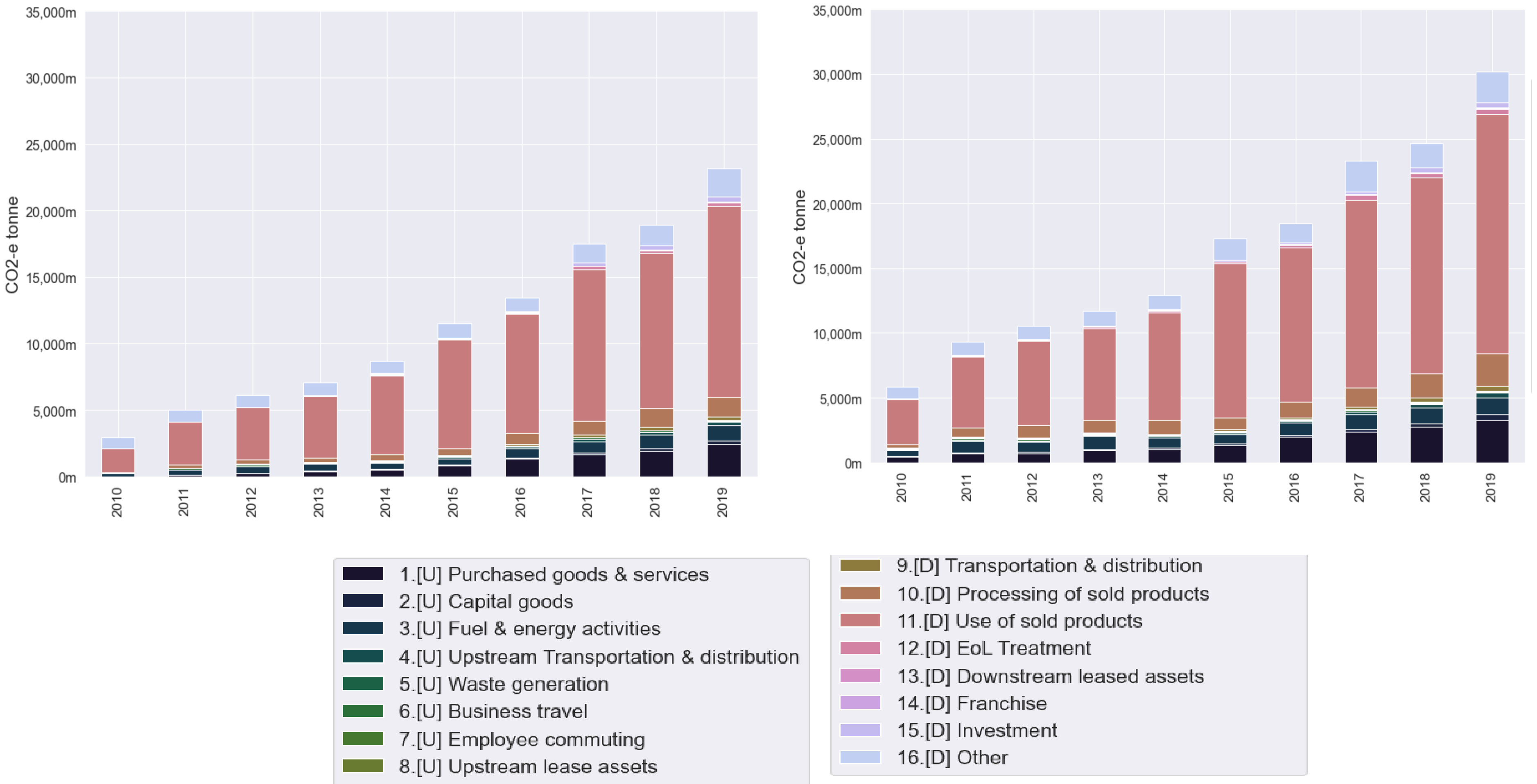


Figure A2 - Emission profiles after filing in the gaps for unreported categories by sector

Notes: This figure summaries the emission profiles before and after filling in the gaps on all unreported categories for the sample of 9,518 observations from 1,972 firms that disclose the composition of Scope 3 in the Bloomberg dataset in 2010-2019 as described in Section 4.2.