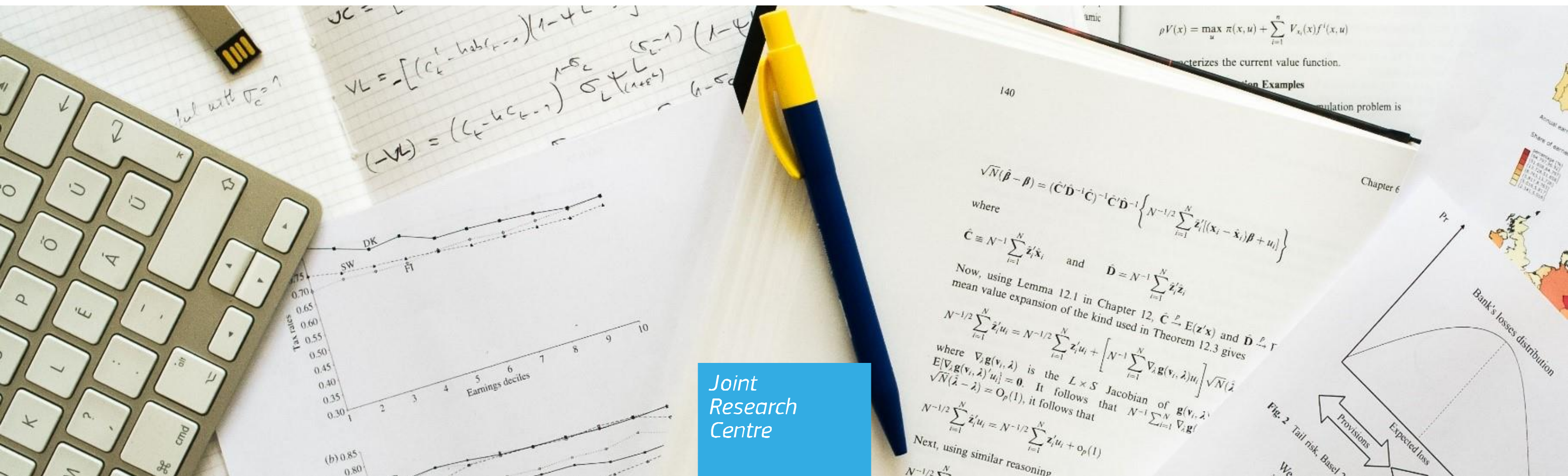
European Commission

External assurance of carbon disclosures indicates possible underestimates in reported European corporate emissions data

Papadopoulos, Georgios

2023

JRC Working Papers in Economics and Finance, 2023/9



Joint Research Centre

Contact information
Name: Georgios Papadopoulos
Address: Joint Research Centre, Via Enrico Fermi 2749, TP xxx, 21027 Ispra (VA), Italy
Email: Georgios.Papadopoulos1@ec.europa.eu and gpapad.gr@gmail.com
Tel.:

EU Science Hub
https://joint-research-centre.ec.europa.eu

JRC 134799

Ispra: European Commission, 2023

How to cite this report: Papadopoulos, G., *External assurance of carbon disclosures indicates possible underestimates in reported European corporate emissions data*, JRC Working Papers in Economics and Finance, 2023/9, European Commission, Ispra, 2023, JRC134799.

Executive summary

The cornerstone of addressing the climate crisis is the reduction of the emitted amount of greenhouse gases. Evidently, it is fundamental to estimate their amounts, at the highest possible quality. At company level, the latter is achieved through a multi-step process known as measurement, reporting, and verification (MRV).

At EU level, several policy initiatives have been taken to regulate the various elements of the MRV procedure for companies such as the EU Emissions Trading System (EU ETS), the Directive 2014/95/EU also known as the non-financial reporting directive (NFRD) and, recently, the Corporate Sustainability Reporting Directive (CSRD). Although these policy initiatives have improved carbon disclosure, recent research suggests that corporate GHG emissions data still suffer from substantial quality issues.

In particular, while an increasing number of companies has been engaging with the first two steps of the MRV process, less than half of them obtained external verification of their sustainability reports, part of which are their carbon disclosures.

Until now, no study investigates the possible existence of systematic differences in reported emissions with respect to their assurance status. This indicates that the profession either ignores or implicitly assumes such differences to be absent or immaterial. This study examines what can be termed as the effect of external assurance on carbon disclosures; i.e. whether there is a systematic difference between reported figures, depending on their verification status and the equally important questions about the direction and magnitude of that difference, in a recent sample of Europe-domiciled firms.

Possible existence of such difference can have significant and interrelated implications for every stakeholder involved. For companies, it has been well documented that carbon performance is positively linked with their financial performance reflecting a lower climate transition, reputation and litigation risk profile. Substantial misstatements in disclosed corporate emissions figures -especially if errors favor a company's carbon performance- could adversely impact all previous types of risk and have serious financial ramifications. Naturally, if these risks materialize, the financial sector will also be impacted. For example, a sizable positive restatement in firms' carbon footprint coupled with a rise in the cost of carbon -either through the introduction of a carbon tax or some other market mechanism- could trigger defaults across firms. This could subsequently spillover to banks exposed to those firms. However, even a benign upward revision of a company's reported emissions could potentially give rise to investment outflows since, as research shows, institutional investors significantly divest from highly emitting firms. Finally, effective design and implementation of policy rests upon a reliable and as accurate as possible, knowledge of what constitutes the very target of such actions; GHG emissions. From assessing the current state, setting reduction goals, choosing and calibrating the instruments to meet them and finally measuring progress towards them.

The analysis clearly identifies a positive effect of external assurance on company Scope 1 emissions, suggesting that average reported direct emissions rise by 4.5% to about 8% when every company engages with external assurance compared to when none does so. However, findings are much weaker in the case of location-based Scope 2 emissions whereas, practically, there weren't robustly detected any differences in reported market-based Scope 2 emissions with respect to their assurance status.

These findings have important policy implications. First, they suggest that companies might underestimate the amount of their direct carbon footprint. While the error's magnitude from a company's perspective might be moderate, on average, it could still be a source of reputational risk and trigger investment outflows. More importantly though, it can have significant ramifications from an economy-wide point of view and the accomplishment of climate neutrality targets. In that regard, it is worth noting that the largest annual reduction in total GHG and energy-related CO_2 emissions at the EU has been recorded in 2020, evidently related to the impact of COVID-19 pandemic measures. The magnitude of these reductions is about 8% and 5%, respectively, comparable to the average effect of assurance on Scope 1 carbon disclosures, but in the opposite direction.

The weaker or even absent effect of assurance on reported indirect emissions from electricity use is also important because it is likely associated with the clear and indisputable nature of their estimation. Based on electricity consumption data and/or contractual information of the electricity purchased, leaves little room for misreporting. Hence, improving clarity and requiring similarly strong evidence for direct emissions would possibly increase their validity.

Once data availability permits, several important extensions of the study could be considered; to assess the effect of the two assurance levels on carbon disclosures; to re-examine the results after CSRD-related disclosures become available; and consider indirect, Scope 3 emissions in the analysis. Intuitively, the hypothesized difference's magnitude should be positively linked with the applied level of assurance. In addition, given the known data quality issues, the effect of assurance should be even more pronounced in Scope 3 emissions.

The most recent evidence indicate that humanity will likely miss the Paris Agreement goal of limiting global warming to well below 2 degrees Celsius and point instead to a 2.4 to 2.8 degrees Celsius temperature rise by the end of the century. It is indicative that at the time of writing of this study (July 2023), the world was experiencing its hottest month ever, breaking several global temperature records. And, alarmingly, scientists warn that this trend will continue, unless immediate action is taken. This study showed that assurance of corporate carbon disclosures could probably provide more satisfactory and certainly more prudent estimates of emitted GHGs by firms. Given the dire climate prospects ahead, it is arguably better to err on the side of caution and act upon it.

Contents

Abstract

Company carbon disclosures are crucial in assessing a firm's impact on the environment, and many policy actions are associated with this information. As a response to the rising demand for transparency and related regulatory requirements, an increasing number of firms discloses information on their greenhouse-gas (GHG) emissions and voluntarily engage with external assurance of the reported information. However, the possible existence of systematic differences in reported emissions with respect to their assurance status is still under-explored. This study investigates the causal effect of third-party assurance on carbon disclosures in a sample of European companies. Findings suggest that non-assuring firms may be under-reporting their direct GHG emissions by up to a magnitude comparable to the largest annual reduction of EU emissions in history. On the contrary, the effect of assurance is much weaker to almost absent in indirect, Scope 2, emissions possibly due to their clear and easily verifiable estimation nature. These findings demonstrate that third-party assurance can provide more reliable and certainly more prudent estimates of corporate GHG emissions which are relevant to corporate sustainability strategy, policymaking and, ultimately, climate change mitigation.

. . . [W]hen you cannot measure it . . . knowledge is of a meagre and unsatisfactory kind.(1)

1 Introduction

Based on the consensus built upon accumulated scientific evidence, policymakers around the world committed to take action to reduce the main cause of global warming, associated with human activity; greenhouse gas (GHG) emissions. Evidently, the cornerstone of such efforts is knowledge of the emitted amount of greenhouse gases, estimated at the highest possible quality.

At company level, the latter is achieved through a multi-step process known as measurement, reporting, and verification (MRV). Over time, an increasing number of companies has been engaging with the first two steps of the MRV process, namely estimating and reporting their emissions, either voluntarily or due to regulatory requirements (KPMG, 2022). However, less than half of them obtained external verification(2) of their sustainability reports, part of which are their carbon disclosures (KPMG, 2022).

Regardless of the underlying reasons behind companies' choice to engage with external verification of their disclosed emissions information, a key question arises: is there a systematic difference between reported figures, depending on their verification status? An affirmative answer, subsequently raises equally important questions about the direction and magnitude of that difference.

Subject to the answers' nature, there can be significant and interrelated implications for every stakeholder involved. For companies, it has been well documented that carbon performance is positively linked with their financial performance reflecting a lower climate transition, reputation and litigation risk profile (Busch and Lewandowski, 2018, Velte et al., 2020, Galama and Scholtens, 2021, and references therein). Substantial misstatements in disclosed corporate emissions figures -especially if errors favor a company's carbon performance- could adversely impact all previous types of risk and have serious financial ramifications(3). Naturally, if these risks materialize, the financial sector will also be impacted. For example, a sizable positive restatement in firms' carbon footprint coupled with a rise in the cost of carbon -either through the introduction of a carbon tax or some other market mechanism- could trigger defaults across firms. This could subsequently spillover to banks exposed to those firms (ECB/ESRB, 2022). However, even a benign upward revision of a company's reported emissions could potentially give rise to investment outflows since, as research shows, institutional investors significantly divest from highly emitting firms (Bolton and Kacperczyk, 2021a). Finally, effective design and implementation of policy rests upon a reliable and as accurate as possible, knowledge of what constitutes the very target of such actions; GHG emissions. From assessing the current state, setting reduction goals, choosing and calibrating the instruments to meet them and finally measuring progress towards them.

This study attempts to answer the aforementioned questions with a focus on Europe-domiciled companies. The EU has consistently been the world's third largest fossil CO_2 emitter based on data from the Emissions Database for Global Atmospheric Research (EDGAR) (Crippa et al., 2022). At the same time, several policy initiatives have been taken at EU level, to regulate the various elements of the MRV procedure for companies. Arguably, some of the most demanding are those associated with the EU Emissions Trading System (EU ETS)(4) which requires firms under its scope to conform with specific, legally binding, rules with respect to monitoring, reporting and verifying their emissions(5). While the disclosed information is, possibly, of the highest quality, it is not an all-encompassing one. By design, it focuses on emissions from individual facilities in certain, energy-intensive, industry sectors(6) and specific greenhouse gases. Research suggests that emissions reported under the EU ETS cover about 40% of reporting companies' total direct emissions (Busch et al., 2022). Another major legislative initiative at the EU level has been the Directive 2014/95/EU also known as the non-financial reporting directive (NFRD)(7). As of 2018, the rules introduced by the NFRD require large,

(1) Sir William Thomson (also known as Lord Kelvin), in a Lecture on Electrical Units of Measurement (3 May 1883), published in Popular Lectures and Addresses Vol. I, p. 73.

(2) *Verification* refers to the process followed to determine whether disclosed information is materially correct. *Assurance* is the outcome of verification, stating that an auditor is confident that the information disclosed is accurate with adequate certainty. Henceforth, these terms will be used interchangeably.

(3) Probably the biggest emissions scandal to date, associated with the manipulation of vehicle emissions by the German automobile firm Volkswagen, did cost the transgressing company 31.3 billion euros in fines and settlements (Reuters, 2020) and had negative reputational spillovers to all German car manufacturers and potentially to the country's image (Aichner et al., 2021).

(4) Directive 2003/87/EC of the European Parliament and of the Council of 13 October 2003 establishing a system for greenhouse gas emission allowance trading within the Union and amending Council Directive 96/61/EC

(5) There are two regulations setting out the rules related to the MRV procedure and its associated processes. The Regulation (EU) 2018/2066 of 19 December 2018 on the monitoring and reporting of greenhouse gas emissions pursuant to Directive 2003/87/EC of the European Parliament and of the Council and amending Commission Regulation (EU) No 601/2012 (Monitoring and Reporting Regulation) and the Regulation (EU) 2018/2067 of 19 December 2018 on the verification of data and on the accreditation of verifiers pursuant to Directive 2003/87/EC of the European Parliament and of the Council (Accreditation and Verification Regulation).

(6) Examples include commercial aviation, oil refineries, steel works, production of iron, aluminium, metals and cement among others.

(7) Directive 2014/95/EU of the European Parliament and of the Council of 22 October 2014 amending Directive 2013/34/EU as regards disclosure of non-financial and diversity information by certain large undertakings and groups. OJ L 330, 15.11.2014, p. 1-9

public interest entities with over 500 employees to publish non-financial information on environmental, social and governance matters. Subsequently, the European Commission has published legally non-binding guidelines on reporting climate-related information, including comprehensive indicators on corporate emissions([8]). The regulation allows Member States to ask for verification of the provided information by an independent assurance services provider. However, an early report indicated that countries largely implemented the minimum required level of auditors' involvement, i.e. to check whether the needed information is present (GRI, CSR Europe and Accountancy Europe, 2017). The latest legislative initiative on sustainability reporting at the EU is the Corporate Sustainability Reporting Directive (CSRD)([9]). It entered into force on January 2023 and applies to 4 times as many companies compared to the NFRD. It requires updated social and environmental information, including mandatory disclosures of companies' direct and indirect emissions. Moreover, it makes assurance of the previous information obligatory. The first reports under the CSRD will be published in 2025 covering financial year 2024. Therefore, until then, carbon disclosure and its assurance will largely remain voluntary with the exception of facilities under the scope of the EU ETS. Although these policy initiatives have improved carbon disclosure, recent research suggests that corporate GHG emissions data still suffer from substantial quality issues (Papadopoulos, 2022).

1.1 Assurance of carbon disclosures

An assurance provider's objective is to express a conclusion about whether a company's GHG statement is not materially incorrect. Different standards detail the exact procedures to be followed, yet they all require some form of evidence collection and calculations reperfomance. In doing so, the assurance provider can catch and correct potential errors thus, improving data quality.

Depending on the amount of evidence obtained and the extent of scrutiny applied, the profession recognizes two levels of assurance, usually termed as limited and reasonable assurance. The former generally requires fewer amount of evidence and a lower level of scrutiny. Consequently, it results in a relatively lower level of confidence regarding the outcome of assurance. The opposite holds for a reasonable assurance engagement.

Two studies document a positive link between assurance and proxies of carbon disclosure comprehensiveness (Dutta and Dutta, 2021, Luo et al., 2023)([10]). Nevertheless, a recurring observation among various, partially overlapping, review papers is that research on carbon disclosure assurance is still limited, despite the fact that the general subject of carbon disclosure has been gaining traction in the academic literature (Hahn et al., 2015, Borghei, 2021, He et al., 2022)([11]). In that growing body of empirical work there are few studies which examine specialized topics such as stakeholders' perceptions with regard to emissions assurance (Green and Li, 2011, Green and Taylor, 2013) or the role of assurer teams' composition on its trust and effectiveness (Kim et al., 2016, Ekasingh et al., 2019). However, the majority tries to identify the firm-level correlates of the choice to obtain external assurance on reported GHG emissions (Zhou et al., 2016, Datt et al., 2018, Datt et al., 2019, Fan et al., 2021) or, more specifically, those associated with the assurance provider's type (accounting vs. non-accounting firms) choice (Green et al., 2017, Datt et al., 2020). Despite study heterogeneity, the literature consistently finds statistically significant links between a small number of factors and the decision to engage with carbon disclosures' external assurance. These include company size, profitability, sector and country-level characteristics.

An important observation follows from this nascent literature on carbon disclosure assurance. To the best of the author's knowledge, no study investigates the possible existence of systematic differences in reported emissions with respect to their assurance status([12]). This indicates that the profession either ignores or implicitly assumes such differences to be absent or immaterial.

1.2 Hypothesis development

The current study hypothesizes that there are systematic differences in reported GHG emissions between companies which externally verify their carbon disclosures and their peers which don't. If existent, they could

([8]) Communication from the Commission - Guidelines on non-financial reporting: Supplement on reporting climate-related information. C/2019/4490, OJ C 209, 20.6.2019, p. 1-30

([9]) Directive (EU) 2022/2464 of the European Parliament and of the Council of 14 December 2022 amending Regulation (EU) No 537/2014, Directive 2004/109/EC, Directive 2006/43/EC and Directive 2013/34/EU, as regards corporate sustainability reporting. PE/35/2022/REV/1, OJ L 322, 16.12.2022, p. 15–80

([10]) Or, at least, completeness of disclosed climate information. Both studies use a carbon disclosure score as a proxy of disclosure quality which measures how comprehensive a firm's answer to CDP's questionnaire has been. See CDP Scoring Methodology for more details (CDP, 2023).

([11]) For an overview of the literature on the broader topic of reporting and assuring sustainability-related non-financial information, the interested reader is referred to the comprehensive studies of (Venter and Van Eck, 2021, Zhou, 2022) and references within.

([12]) The study of (Michelon et al., 2019) examines the link between assurance and restatements in sustainability reports. Although it shares some similarities, it has different focal points both in scope (sustainability report) and variable considered (restatements).

manifest at an aggregate level in two possible directions[13]: companies with externally verified carbon disclosures reporting, on average, higher emissions compared to their non-verified peers or the opposite. Any of them could be the result of non-mutually exclusive, systematic actions by the two parties involved; reporting companies and assurance providers.

1.2.1 The direction of the difference

A negative, on average, difference in reported emissions between companies with and those without verified carbon disclosures would indicate that the latter overestimate their emissions and/or the former underestimate them. However, these are two rather remote possibilities. On the one hand, non-verifying companies do not have any incentive to purposely over-report their emissions, thus exhibiting a more polluting profile and higher climate transition risk. On the other hand, systematic under-reporting by verifying firms would raise concerns of greenwashing and misconduct for both the benefiting firms and their assurance providers. In case of revealed malpractice both, but especially the latter, would face substantial reputational damages. Therefore it is unlikely that any of these mechanisms is at play in a consistent and large scale manner. A more plausible alternative yielding this outcome is a form of selection bias; companies who made a considerable effort to reduce their GHG emissions might seek assurance to demonstrate their better performance. However, research shows that assurance gets the lion's share of MRV costs, ranging from 40% to up to 80% of the total (Bellassen et al., 2015). Hence, it might be difficult to justify the existence of this mechanism, unless benefits from external assurance outweigh its substantial costs.

On the contrary, there are two conceivable alternatives resulting in a positive average difference in reported emissions between verifying and non-verifying companies. As the level of scrutiny is low and the likelihood of outside users uncovering errors is very limited, non-verifying companies may perceive that risks connected with misreporting are insignificant. Hence, they might be tempted to err on the side which favors them thus under-reporting their emissions. A more subtle, potentially concurrent, mechanism acting in the same direction is associated with the assurance providers. In particular, it is related to a potential attempt of theirs to increase the perceived value of their services through the provision of more conservative (i.e. higher) emissions estimates. A relevant behavior has been documented in the study of (Michelon et al., 2019) where the authors find a positive link between restatements in sustainability reports and assurance. They argue that this is a strategy by assurance providers which allows them to gain market share and their clients to exhibit an image of transparency.

This study states the following formal hypothesis with respect to the difference's direction:

H1: The difference in reported emissions between companies which engage with external assurance of their carbon disclosures and their peers which don't is positive, on average.

1.2.2 The magnitude of the difference

The magnitude of the hypothesized difference is arguably irrespective of the mechanism behind it. Any of the two would suggest a small- to medium-sized systematic difference between the verifying and non-verifying companies, though for different reasons. A plausibility check could detect unjustifiably low emissions by the non-verifying and presumably under-reporting firms and thus draw attention to them. Hence, should they go in that direction, they would rather report lower, yet reasonably close figures to their peers. Furthermore, the, supposedly strategically acting, assurance providers might face a potential conflict of interest. By being overly conservative they would harm their client's environmental profile. Consequently, they would possibly refrain from reporting substantially higher emissions compared to the non-verifying companies.

Hence, the second hypothesis is the following:

H2: The hypothesized difference's magnitude is moderate, at most.

The two hypotheses described above constitute what can be called "*the effect of assurance on carbon disclosures*". Finding supporting evidence of an effect of assurance on carbon disclosures is important despite the fact that this study does not identify the exact underlying mechanism. The most recent evidence indicate that humanity will likely miss the Paris Agreement goal of limiting global warming to well below 2°C and point instead to a 2.4°C to 2.8°C temperature rise by the end of the century (UNEP, 2022). Therefore, it is crucial if external assurance helps correct misstatements related to systematic under-reporting by non-verifying companies. Even if assurance providers' internal motives are behind verifying companies' moderate emissions over-reporting, it is arguably better to err on the side of caution, given the aforementioned dire predictions.

(13) At firm level, differences could occur as a result of innocent mistakes due to the complexities involved with emissions estimation and companies' potential lack of expertise. However, in that case they would appear in both directions, e.g. over-reporting as well as under-reporting emissions data, thus canceling out on average.

2 Data

The study investigates econometrically the causal effect of assurance on company-reported Scope 1 and on Scope 2 emissions. The former are direct emissions from owned or controlled sources while the latter are emissions from the generation of purchased electricity used by the company (WRI and WBCSD, 2004). Scope 2 emissions can be estimated using two methods. The first method, termed location-based, estimates emissions according to the average emissions intensity of grids on which energy consumption occurs. The second method, called market-based, uses contractual information, such as renewable energy certificates, on the electricity that companies have purposefully chosen to estimate their Scope 2 emissions([14]). Unfortunately, reported data on indirect emissions from activities along the supply chain (Scope 3) and their assurance are fragmented to a degree which does not allow for their consistent analysis.

The sample consists of Europe-domiciled([15]) firms from 2017 until 2021. It contains information on, practically, every company publicly reporting emissions data. One limitation of this, and every firm-level, emissions dataset is that it mostly focuses on large, listed companies whereas representation of SMEs is minimal. This is due to the fact that the latter are not required to and usually don't disclose such information. Therefore, the results are applicable to this specific firm population. However, they are still highly relevant since, arguably, large firms are the most polluting ones.

The year 2017 is a natural origin because then the Task Force on Climate-Related Financial Disclosures (TCFD)([16]) published its final report with recommendations for climate-related disclosures. Prior to that, the further in the past, the scarcer information on corporate emissions and their assurance status is([17]).

2.1 Corporate emissions

2.1.1 Overview

Reported corporate emissions data are from ICE Climate Transition Finance (ICE CTF)([18]) and are expressed in tons of CO_2 equivalent. They include information on their assurance such as its level, the standard applied and the share of total reported emissions assured. In most cases assurance covers above 95% of companies' global emissions. The analysis focuses on the latter in order to capture the effect of assurance on reported emissions while minimizing the influence of low coverage([19]).

Table 1 presents the descriptive statistics of the log-transformed emissions variables of interest.

([14]) Greenhouse Gas Protocol, GHG Protocol Scope 2 Guidance. An amendment to the GHG Protocol Corporate Standard.

([15]) The companies analyzed in the study are domiciled in the following countries: Austria, Belgium, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Lithuania, Luxembourg, the Netherlands, Norway, Poland, Portugal, Romania, Spain, Sweden, Switzerland and the UK.

([16]) The TCFD has been established in December 2015 by the Financial Stability Board and tasked to develop recommendations on the type of company disclosures that would support financial market participants appropriately assess and price climate change related risk.

([17]) E.g. the study of (Papadopoulos, 2022) indicates that, in the EU, reported emissions data in 2008 are 5 to 7 times fewer than in 2020, depending on data provider.

([18]) Previously known as Urgentem. More information: https://www.ice.com/data-services/sustainable-finance-data/climate-data [Accessed: August 2023]

([19]) The Annex provides more details about the low-coverage observations.

Table 1: Descriptive statistics of corporate emissions data.

Statistic	$LOGS1$	$LOGS2_{loc}$	$LOGS2_{mkt}$
Minimum	-2.30	-1.39	-0.26
Maximum	19.00	17.33	17.18
Mean	10.00	10.08	9.80
Median	9.91	10.20	10.00
Std. Dev.	3.24	2.61	2.91
Kurtosis	3.07	3.19	2.92
Skewness	0.06	-0.33	-0.36
Observations	4913	4815	2867

Note: All variables are transformed by the natural logarithm of the respective reported corporate emissions data. $LOGS1$ refers to Scope 1, $LOGS2_{loc}$ to location-based Scope 2 and $LOGS2_{mkt}$ to market-based Scope 2 emissions.
Source: ICE CTF and author's calculations.

The natural logarithms of every emissions variable exhibit close means and medians, kurtoses around 3 and absolute skewnesses below 1. Overall, these observations suggest that the data are symmetrically distributed and, in particular, approximately normally so (Hair et al., 2010). Scope 1 and location-based Scope 2 emissions data are equally well populated with almost 5000 observations over the sample period. On the contrary, market-based Scope 2 emissions data have about 40% less observations than the other two variables. This is probably related to the nature of that type of data. Scope 1 emissions and location-based Scope 2 emissions are likely easier to estimate. However, market-based Scope 2 ones require that companies have purposely chosen to purchase renewable energy and have the contractual information to prove it. Thus, although reported figures might be more accurate than the location-based, many companies perhaps did not actively choose that alternative.

These observations are neither evenly spread over time nor across the different levels of assurance. Figures 2 and 3 present the distributions of emissions data in time and by level of assurance.

Figure 1: Scope 1 emissions' sample size by year and level of assurance.

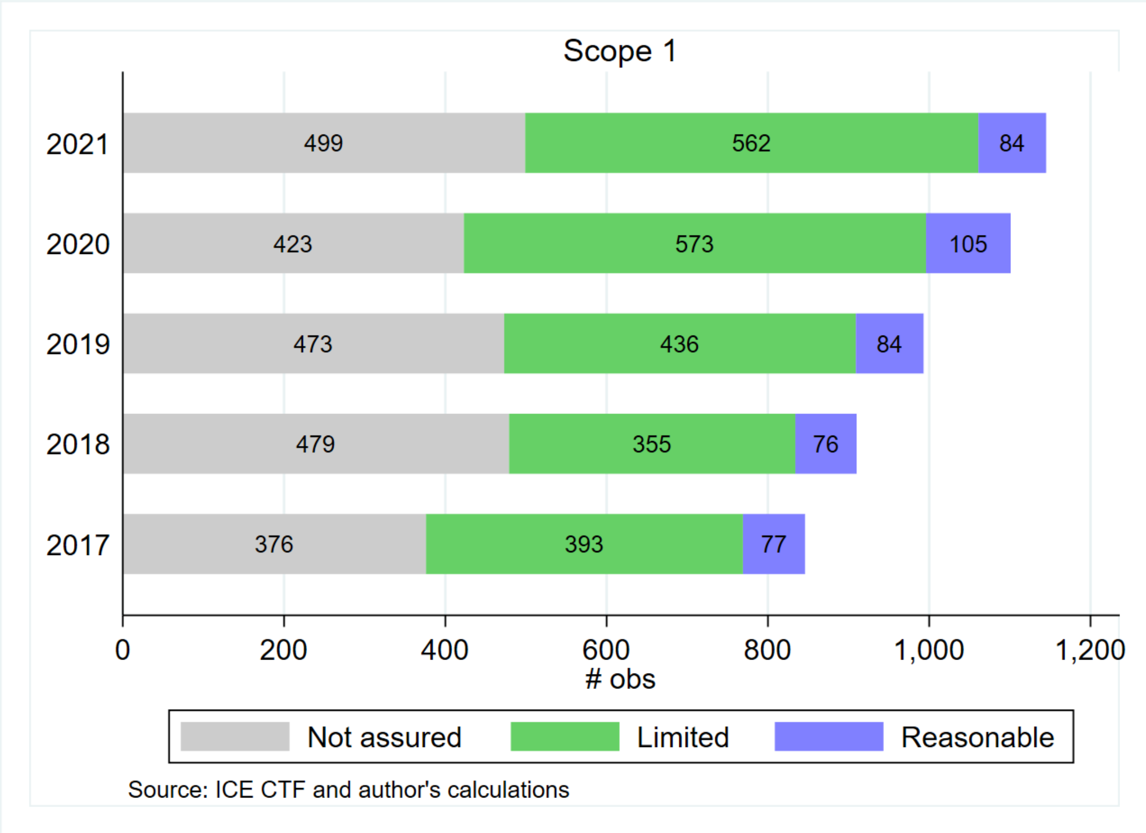
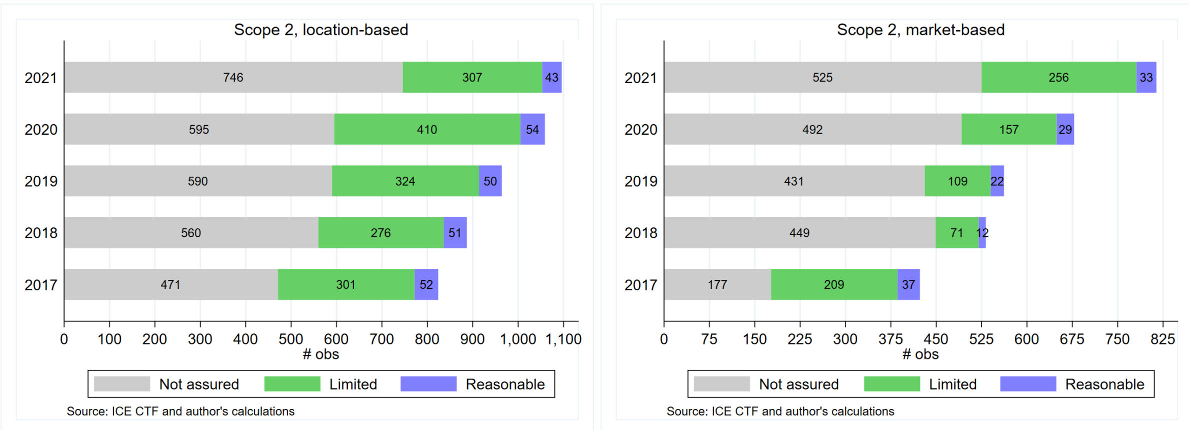


Figure 2: Scope 2 emissions' sample size by year and level of assurance.



In each Scope 2 estimation method the non-assured status applies if emissions data are disclosed but assurance information for that method is missing.

Some general patterns emerge from Figures 1 and 2. First, emissions disclosures show a steady rise over time. Scope 1 and location-based Scope 2 sample sizes increase by more than 33% over the dataset's time range while market-based Scope 2 emissions data availability almost doubles in the same period.

With respect to assurance, its highest level has always the fewest observations regardless of emissions Scope or year. Finally, another common pattern, especially between Scope 1 and location-based Scope 2 emissions, is that the shares of each assurance level -including not assured- remain roughly stable over time. However, in Scope 1 emissions, assured information usually exceeds not assured in availability, at times even considering only the lowest assurance level. This is not the case in Scope 2 emissions, where the number of observations without assurance is always higher than the assured ones. A notable exception is the first year in the sample of market-based Scope 2 emissions. Nevertheless, in subsequent years sample sizes of assured emissions information are a small, yet increasing, fraction of the total disclosed figures.

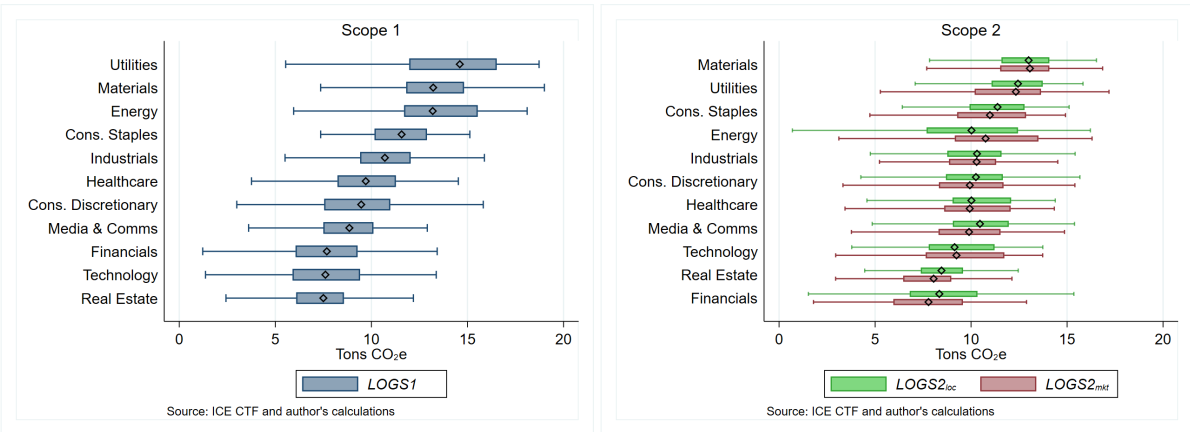
It is worth noting that characterizing Scope 2 emissions as non-assured can be done in various ways due to their two possible estimation methods. The conventional way of considering an observation as non-assured is based on the absence of associated assurance information. This can naturally apply on data from each of the two Scope 2 emissions estimation methods. However, companies often report both location- and market-based Scope 2 emissions but provide assurance data only for one of the two. In that case, a less constrained alternative would be to assume that assurance "extends" to both of them. Thus, the non-assured status would apply to Scope 2 data without assurance information in both estimation methods. A more conservative variation of the previous would be to neither extend the assurance status to observations that it doesn't explicitly refer to, nor classify the latter as non-assured. Rather, exclude from the analyses Scope 2 observations with absent assurance information whose estimation method counterpart is assured. This study applies the conventional way described above for assigning observations' assurance status.

2.1.2 Sectoral patterns of emissions and their assurance

The dataset contains information on firms' sectoral classification which can reveal interesting aspects of emissions and assurance heterogeneity. The classification scheme is designed by ICE CTF and is similar to the Global Industry Classification Standard (GICS). At the highest level of grouping it includes 11 sectors; utilities, materials, energy, consumer staples, industrials, healthcare, consumer discretionary, media and communications, financials, technology and, finally, real estate.

Figure 3 presents the medians and dispersion statistics of emissions' distributions by sector.

Figure 3: Descriptive statistics of sectoral emissions.

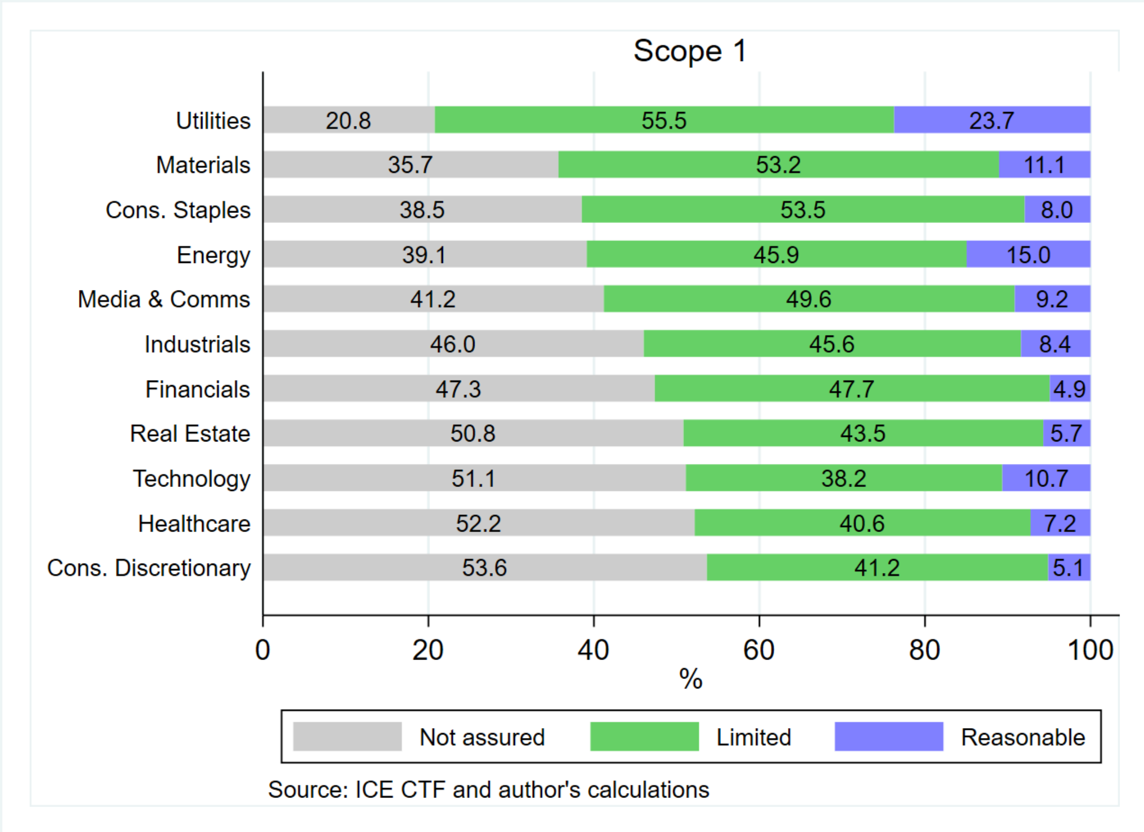


Diamonds denote the median while the lower/upper hinges of each box mark the first/third quartiles of variables' distributions. Whiskers are placed beyond the first/third quartiles by 1.5 times of each distribution's interquartile range.

As expected, corporate emissions exhibit noticeable heterogeneity across sectors. With regard to direct, Scope 1, emissions the three most polluting sectors are utilities, materials and energy. In addition to high medians and lower/upper quartiles, they have high dispersion as shown by the associated whiskers. In the subsequent five sectors most order statistics show a gradual decline until the three less emitting ones -financials, technology and real estate- which have similar emissions distributions. The most and least emitting sectors maintain their status in Scope 2 emissions. Materials, utilities and energy are at the top with the latter practically sharing its position with the consumer staples sector. Similarly, technology, real estate and financials are the sectors with the lowest indirect emissions, overall.

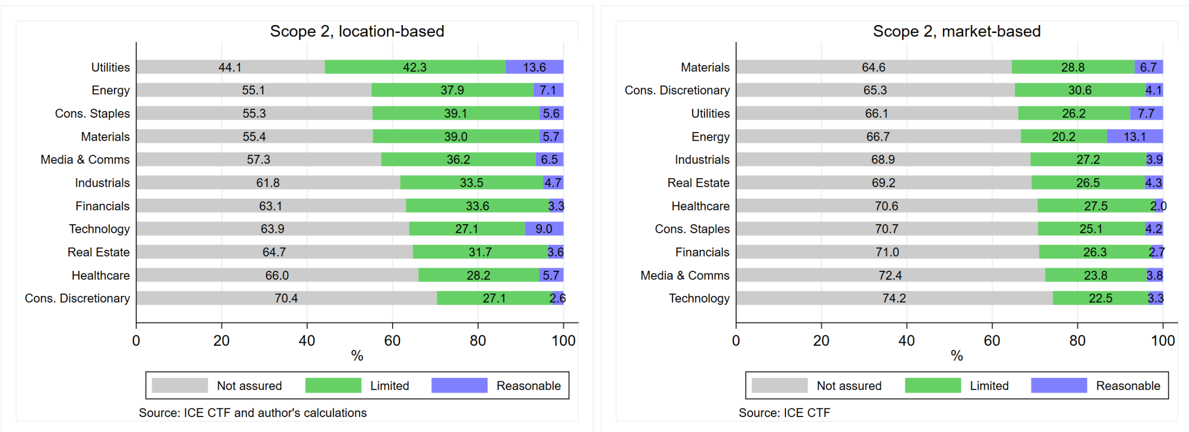
Another relevant element for the subsequent analysis is the potential diversity in assurance engagement across sectors. Figure 4 and Figure 5 show the shares of reported emissions data with reasonable, limited and no assurance within each sector.

Figure 4: Scope 1 emissions assurance engagement by sector and level of assurance.



The horizontal axis shows the share of Scope 1 emissions reports by their assurance status.

Figure 5: Scope 2 emissions assurance engagement by sector and level of assurance.



The horizontal axis shows the share of Scope 2 emissions reports by their assurance status.

The graphs in Figures 4 and 5 reveal similar heterogeneity in assurance engagement across sectors with the one observed in emissions levels. In particular, the most polluting sectors have lower shares of non-assured reports in any emissions Scope compared to the less polluting ones. Between direct and indirect emissions, the former are assured more often than the latter. Focusing on the two Scope 2 estimation methods, market-based one is the least frequently assured which is puzzling given that the information accompanying its disclosure should make it more accessible to assurance.

Possibly there is an association between a sector's emissions levels and assurance engagement. The more polluting an industry is, the more exposure to the public scrutiny it gets and the more focus of policy actions to reduce its emissions it attracts. A combination of the two, along with potential behavioral spillover effects within each sector might be the reasons behind these patterns. The Annex includes a more elaborate discussion on the matter.

2.2 Control variables

In addition to emissions data, the empirical part uses balance sheet information sourced from Refinitiv Eikon, a widely used commercial data provider (see Table 3 in the Annex). These control for factors that research has linked with corporate GHG emissions and the decision to engage with external assurance.

Despite study-specific findings, there is a consensus in the literature about the statistically significant correlates of company GHG emissions and the decision to externally assure carbon disclosures (see Table 2 in the Annex). Interestingly, many of them such as company size, profitability, capital expenditures, the stock of physical capital and leverage are associated to both company emissions and disclosures' external assurance. In addition to balance sheet variables, firm sectoral classification and country-level characteristics are consistently among the important determinants of both variables of interest.

Firm size is positively associated with the level of emissions and carbon disclosure assurance. Larger firms should, generally, have higher emissions than smaller firms (Griffin et al., 2017, Haque, 2017, Hassan and Romilly, 2018, Alam et al., 2019, Bolton and Kacperczyk, 2021a, Bolton and Kacperczyk, 2021b). Moreover, larger companies possibly attract more media and stakeholder scrutiny of their environmental practices, hence they would tend to engage more with assurance to signal their environmental responsibility (Zhou et al., 2016, Datt et al., 2018, Datt et al., 2019, Fan et al., 2021). Similarly, more profitable firms should be more able to cover the additional costs for assurance (Zhou et al., 2016, Datt et al., 2018, Fan et al., 2021). The literature associates capital expenditure intensity with the purchase of new and potentially less polluting equipment (Haque, 2017). This association yields a negative link with absolute emissions and a positive one with the decision to assure carbon disclosures (Bolton and Kacperczyk, 2021a, Bolton and Kacperczyk, 2021b, Fan et al., 2021). Another statistically significant determinant of corporate emissions and assurance engagement is the stock of physical capital (Haque, 2017, Bolton and Kacperczyk, 2021a, Bolton and Kacperczyk, 2021b, Zhou et al., 2016). Also, studies empirically link firm leverage both with GHG emissions (Alam et al., 2019, Bolton and Kacperczyk, 2021a, Bolton and Kacperczyk, 2021b) and carbon disclosure assurance (Datt et al., 2018). Interestingly, leverage could be related to assurance in both directions. It could act as a disciplining device, motivating highly leveraged firms to have more transparent and credible disclosures to reduce information asymmetries with their lenders. On the other hand, high leverage could prevent companies from committing financial resources to purchase assurance services. Finally, all aforementioned studies include sector- and country-level information to account for differences among industries or country-specific policies.

3 Results

To examine the effect of assurance on carbon disclosures, this study uses the potential-outcomes framework. In particular, it estimates the average treatment effect (ATE) which measures the average difference in reported emissions between companies which externally assure their disclosures and those which don't.

Among the various available estimators, the analysis employs the inverse-probability-weighted regression-adjustment estimator (IPWRA). The estimator first models the probability of a company engaging with external assurance (i.e. treatment assignment). Then, uses the estimated probabilities as weights to model reported corporate emissions (outcome). The IPWRA estimator has the advantage of being doubly robust, which means that it remains consistent even if one of the two models is misspecified, provided that the other is correctly specified. For a more detailed description, the reader can refer to (Wooldridge, 2010).

Both models are functions of firm characteristics. In particular, the literature recommends balancing confounding covariates, associated with both treatment and outcomes, instead of variables that only affect the treatment-selection process to achieve lower bias and variance of the treatment-effect estimate (Austin and Stuart, 2015). Therefore, the company-level controls used in the models are those which research has identified as important determinants of both company emissions and the decision to engage with external assurance. Table 2 describes these variables and provides information on the studies which associate them with the dependent variables of interest.

The first step uses a probit model of the following form:

$$Pr(D_{i,t} = 1|X_{i,t}) = \Phi(aX_{i,t}),$$

where $D_{i,t}$ is a dummy variable distinguishing companies with assured carbon disclosures ($D_{i,t} = 1$) and without ($D_{i,t} = 0$) and $X_{i,t}$ a vector of company-level controls.

The next stage uses the predicted inverse probabilities as weights into a linear regression model of corporate emissions and firm-level characteristics:

$$Y_{i,t} = \beta D_{i,t} + aX'_{i,t} + \epsilon_{i,t},$$

where $Y_{i,t}$ refers to each of the emissions Scope examined in this study ($Y_{i,t} = \{LOGS1, LOGS2_{loc}, LOGS2_{mkt}\}$) and $X'_{i,t}$ a vector of control variables which does not necessarily need to be the same as above. Standard errors are always clustered at company level.

The control variables included in the first-stage, treatment assignment, model are the logarithm of total assets (*LOGA*), the ratio of yearly revenues to total assets (*ROA*), the ratio of capital expenditures to total assets (*CAPXA*), the logarithm of gross plant, property and equipment (*LOGPPE*) and the ratio of total debt to total assets (*LEV*). A sector-specific dummy (*EMINS*) has been also used. It assumes the value of 1 if a company belongs to an emissions intensive sector such as utilities, materials or energy and is 0 otherwise. The subsequent outcome model is a function of the previous company-level variables with the addition of a country dummy([20]). The rationale of using the latter only in the regression adjustment model is due to the few observations per country which limit the number of statistical matches. Therefore, in the trade-off between matching by sector and country, the former is preferred. This is because differences are likely to be more significant at a sector level rather than a country level due to the relatively uniform policy environment across the EU. Despite that, the potential differences between countries are taken into account by adding the relevant variable in the outcome model.

The analysis is executed separately for each cross-section, i.e. $t = \{2017, 2018, 2019, 2020, 2021\}$, with both the level of reported emissions and their assurance status measured at the same time. Reverse-causality issues are unlikely to exist because of the specific setup of measuring and reporting relevant data. Firms report emissions at the end of each fiscal year, as every balance sheet variable, and engage with carbon disclosure assurance before emissions figures are reported. Consequently, treatment assignment should be independent of potential outcomes after accounting for relevant covariates.

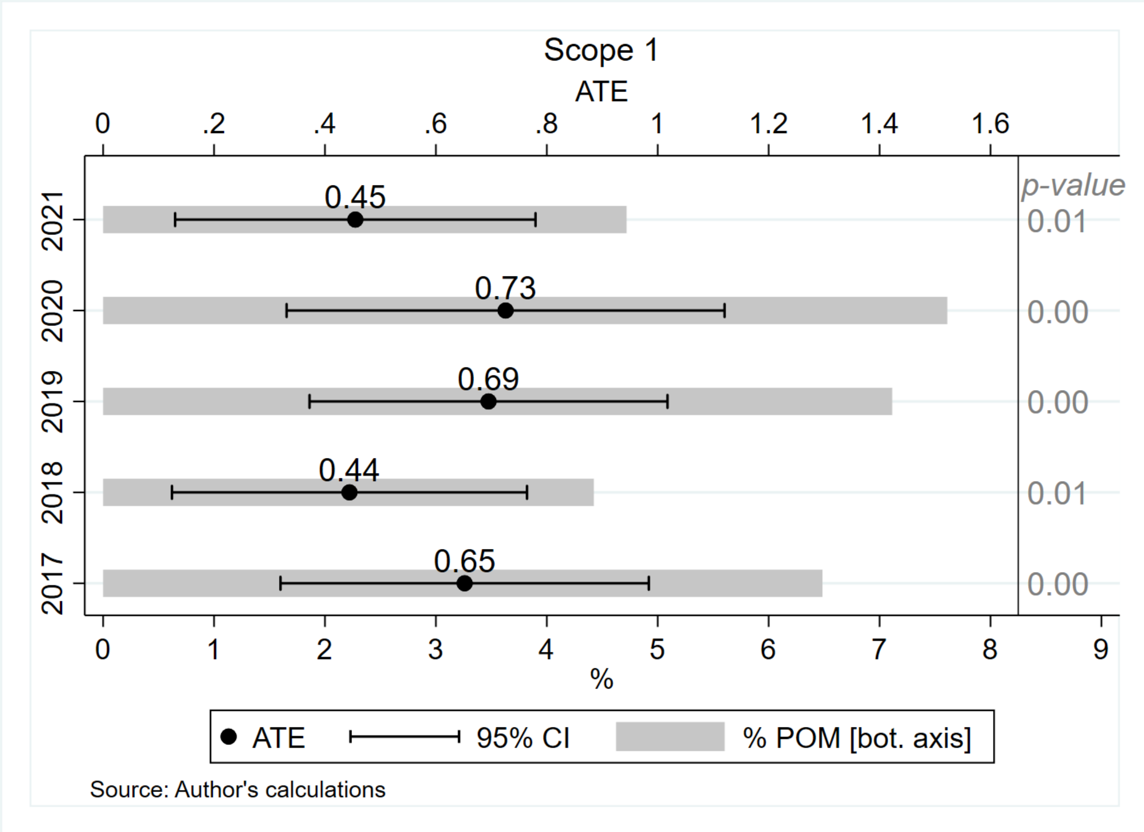
The study's Annex contains model diagnostics and additional analyses that assess the robustness of the results. These include ATE estimates from additional estimators as well as endogeneity tests to examine the presence of unobservable factors that affect both treatment assignment and the outcome. The results of these analyses strengthen the validity of the main findings and demonstrate that they are not dependent on particular modeling assumptions.

3.1 Scope 1 emissions

Figure 6 shows the estimated ATEs, their statistical significance and an intuitive expression of their magnitude, as a percentage of the untreated potential-outcome means, for each year.

([20]) The only exception is the analysis of market-based Scope 2 emissions for year 2020 in which country has been included in both treatment and outcome models to address estimation convergence issues.

Figure 6: Average treatment effect of external assurance on reported Scope 1 emissions.



Several findings emerge from Figure 6. First, ATEs are always positive and statistically significant as their 95% confidence intervals and the respective p-values suggest. This demonstrates the validity of hypothesis *H1* about the existence and the direction of an effect of assurance on reported Scope 1 emissions.

The magnitude of the ATE is more intuitively expressed as a share of the untreated potential-outcome means. As shown by the gray bars in Figure 6, average reported corporate emissions rise by an estimated 4.5% to almost 8% when every company assures their emissions disclosures compared to what it would have been if none did so. These numbers support hypothesis *H2*, indicating an assurance effect of low- to medium-magnitude, from a company's perspective. However, from an economy-wide view, this is a very substantial difference. According to data from the European Environment Agency, the largest yearly reductions in total emissions in the EU-27 since 1990 have also the same magnitude; approximately 8%. Specifically, these happened during two distinct periods: in 2009 following the Global Financial Crisis, and in 2020 during the first year of the COVID-19 pandemic.

Notice that ATE's magnitude fluctuates over time, without a systematic pattern. This can be due to sample differences through the years and/or changes in reporting/assurance practices from the companies analyzed and their assurance providers.

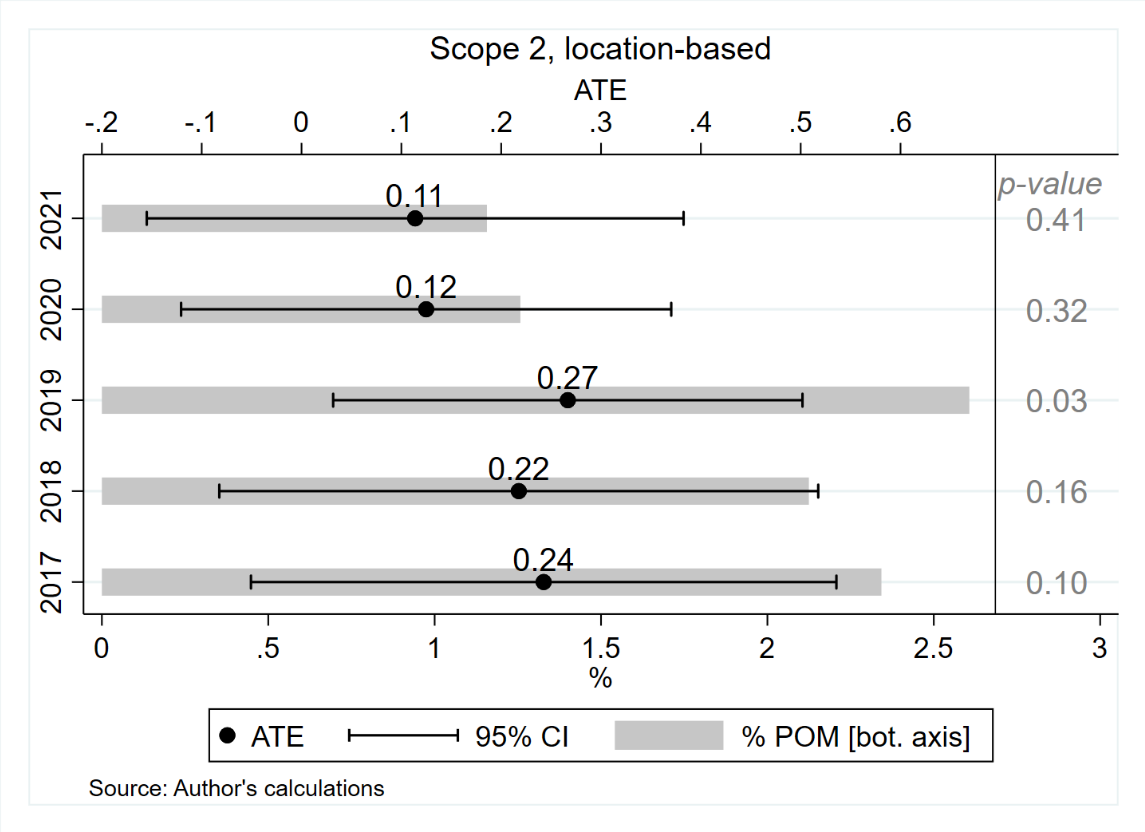
3.2 Scope 2 emissions

Following the division between the two estimation methods for Scope 2 emissions, Figures 7 and 8 present the results from the respective analyses below.

3.2.1 Location-based approach

The findings presented in Figure 7 suggest that external assurance has little or no impact on reported location-based Scope 2 emissions.

Figure 7: Average treatment effect of external assurance on reported location-based Scope 2 emissions.



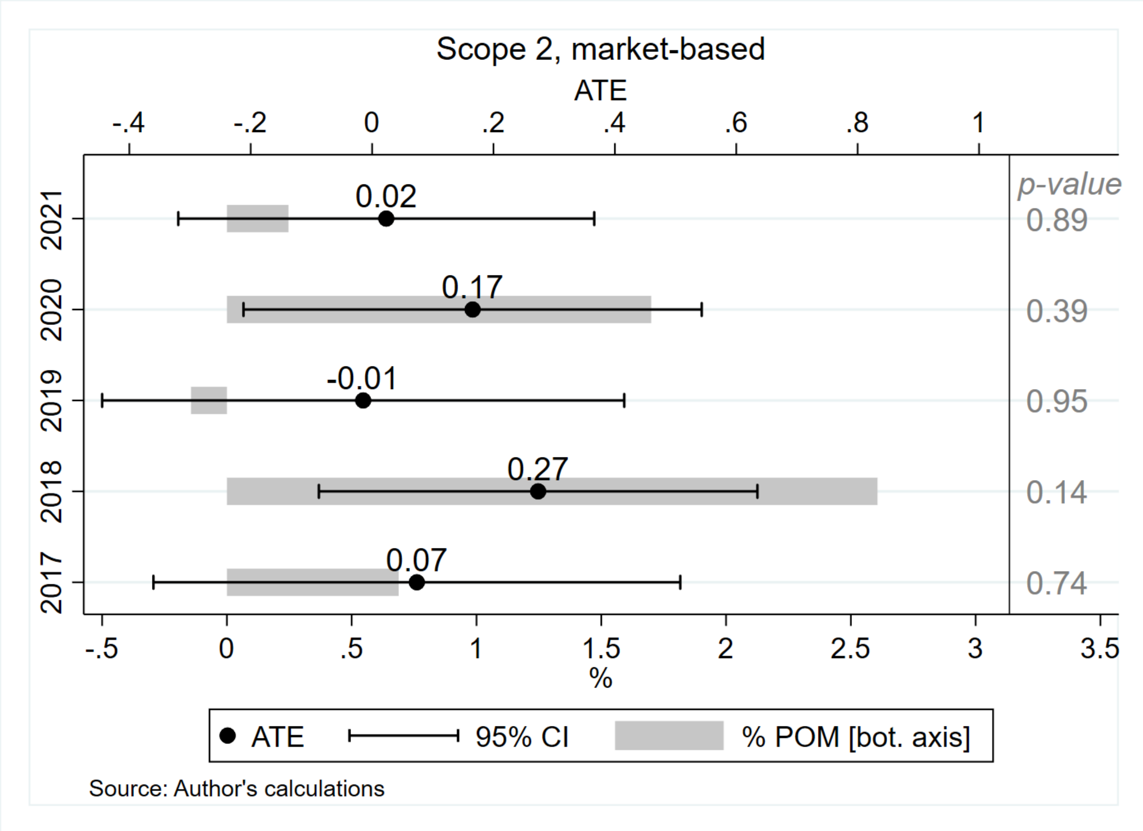
Results in Figure 7 show mixed evidence on the existence of systematic differences in location-based Scope 2 emissions between companies who verify their reports and those who don't. In the two most recent years estimated ATEs are clearly statistically insignificant as the wide confidence intervals and associated p-values indicate. In 2017 and 2019, where ATEs are statistically significant, the magnitude of the effect is considerably smaller compared to the one in Scope 1 emissions. In particular, location-based Scope 2 emissions would be around 2.5% higher, on average, when every company assures their disclosures relative to the case when none does so.

This pattern could be due to the rigid nature of Scope 2 emissions estimation, based on the consumption of electricity, heating and cooling multiplied by the relevant emission factor of the utility grid in which the company operates. The latter depends on the mix of energy sources used by the suppliers to generate the electricity being consumed. These elements are possibly easier to collect reliable information for and therefore minimize any sources of differences between companies with and without external assurance.

3.2.2 Market-based approach

The results in Figure 8 reveal a generally null or, at most, very weak effect of assurance on market-based Scope 2 emissions information.

Figure 8: Average treatment effect of external assurance on reported market-based Scope 2 emissions.



After controlling for confounding variables, there is no discernible difference between externally assured market-based Scope 2 emissions and non-assured ones, in any year. In light of the previous results, this is an expected pattern since estimation of market-based Scope 2 emissions requires concrete proof in the form of contractual information for purchased electricity leaving, virtually, no room for misreporting. This likely explains the absence of systematic differences in reported market-based Scope 2 emissions figures regardless of their assurance status.

4 Conclusions

This study examined a largely overlooked topic by the literature; the existence, direction and magnitude of systematic differences in reported corporate GHG emissions, between companies who externally assure their disclosures and their peers who don't, in a recent sample of Europe-domiciled firms. The two hypotheses, constituting what is termed as the effect of assurance on carbon disclosures, are that such systematic differences exist in favor of the non-assuring companies, while their magnitude is moderate, at most.

The analysis has clearly identified a positive effect of external assurance on company Scope 1 emissions, suggesting that average reported direct emissions rise by 4.5% to about 8% when every company engages with external assurance compared to when none does so. However, findings are much weaker in the case of location-based Scope 2 emissions whereas, practically, there weren't robustly detected any differences in reported market-based Scope 2 emissions with respect to their assurance status.

Beyond the fact that this study is the first to document an effect of assurance on reported carbon emissions, its findings have important policy implications. First, they suggest that companies might underestimate the amount of their direct carbon footprint. While the error's magnitude from a company's perspective might be moderate, on average, it could still be a source of reputational risk and trigger investment outflows. More importantly though, it can have significant ramifications from an economy-wide point of view and the accomplishment of climate neutrality targets. In that regard, it is worth noting that the largest annual reduction in total GHG([21]) and energy-related (IEA, 2022) CO_2 emissions at the EU has been recorded in 2020, evidently related to the impact of COVID-19 pandemic measures. The magnitude of these reductions is about 8% and 5%, respectively, comparable to the average effect of assurance on Scope 1 carbon disclosures, but in the opposite direction.

([21]) According to data by the UNFCCC GHG Data Interface: https://unfccc.int/topics/mitigation/resources/registry-and-data/ghg-data-from-unfccc [Accessed: July 2023]

The weaker or even absent effect of assurance on reported indirect emissions from electricity use is also important because it is likely associated with the clear and indisputable nature of their estimation. Based on electricity consumption data and/or contractual information of the electricity purchased, leaves little room for misreporting. Hence, improving clarity and requiring similarly strong evidence for direct emissions would possibly increase their validity.

These results are important, despite the fact that this study does not identify the underlying mechanism which generates the identified patterns. These can be either due to under-reporting by companies or by assurers who act in a strategic manner, trying to increase the perceived value of their services, or both. The implementation of the CSRD and its assurance requirement might alter the behavior of the latter. A future study could investigate if a consistent pattern emerges after the full implementation of the CSRD and determine if it can be linked to changes in the behavior of external assurers. One limitation of the analysis lies with the sample which covers mostly large, listed companies. This is a general feature of every firm-level emissions dataset and makes the results specific to this company universe. Nevertheless, findings are highly relevant since these companies are likely the most polluting ones. The CSRD requires also listed SMEs to report on sustainability thus, when associated data become available, a future study could re-examine these results and further test their robustness.

Finally, once data availability permits, two important extensions of the study would be to assess the effect of the two assurance levels on carbon disclosures and consider indirect, Scope 3 emissions in the analysis. Intuitively, the hypothesized difference's magnitude should be positively linked with the applied level of assurance. Also, given the known data quality issues, the effect of assurance should be even more pronounced in Scope 3 emissions.

At the time of writing of this study, the world was experiencing its hottest month ever, breaking several global temperature records. And, alarmingly, scientists warn that this trend will continue, unless immediate action is taken ([22]). This study showed that assurance of corporate carbon disclosures could probably provide more satisfactory and certainly more prudent estimates of emitted GHGs by firms. Given the ominous climate prospects ahead, it is arguably better to err on the side of caution and act upon it.

([22]) Copernicus Climate Change Service (C3S), July 2023 sees multiple global temperature records broken, URL: https://climate.copernicus.eu/july-2023-sees-multiple-global-temperature-records-broken [Accessed: 28 July 2023].

References

AccountAbility, 'AA1000 Assurance Standard v3', Tech. rep., 2020. URL `https://www.accountability.org/static/3ff15429033873cdc775212ca63572fb/aa1000as_v3_final.pdf`.

Ackers, B. and Eccles, N. S., 'Mandatory corporate social responsibility assurance practices: The case of King III in South Africa', *Accounting, Auditing & Accountability Journal*, 2015.

Aichner, T., Coletti, P., Jacob, F. and Wilken, R., 'Did the Volkswagen emissions scandal harm the "made in Germany" image? a cross-cultural, cross-products, cross-time study', *Corporate Reputation Review*, Vol. 24, 2021, pp. 179–190.

Alam, M. S., Atif, M., Chien-Chi, C. and Soytaş, U., 'Does corporate R&D investment affect firm environmental performance? Evidence from G-6 countries', *Energy Economics*, Vol. 78, 2019, pp. 401–411.

Austin, P. C., 'Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples', *Statistics in medicine*, Vol. 28, No 25, 2009, pp. 3083–3107.

Austin, P. C., 'An introduction to propensity score methods for reducing the effects of confounding in observational studies', *Multivariate behavioral research*, Vol. 46, No 3, 2011, pp. 399–424.

Austin, P. C. and Stuart, E. A., 'Moving towards best practice when using inverse probability of treatment weighting (iptw) using the propensity score to estimate causal treatment effects in observational studies', *Statistics in medicine*, Vol. 34, No 28, 2015, pp. 3661–3679.

Bellassen, V., Stephan, N., Afriat, M., Alberola, E., Barker, A., Chang, J.-P., Chiquet, C., Cochran, I., Deheza, M., Dimopoulos, C. et al., 'Monitoring, reporting and verifying emissions in the climate economy', *Nature Climate Change*, Vol. 5, No 4, 2015, pp. 319–328.

Bolton, P. and Kacperczyk, M., 'Do investors care about carbon risk?', *Journal of financial economics*, Vol. 142, No 2, 2021a, pp. 517–549.

Bolton, P. and Kacperczyk, M., 'Global pricing of carbon-transition risk', 2021b.

Borghei, Z., 'Carbon disclosure: A systematic literature review', *Accounting & Finance*, Vol. 61, No 4, 2021, pp. 5255–5280.

Busch, T., Johnson, M. and Pioch, T., 'Corporate carbon performance data: Quo vadis?', *Journal of Industrial Ecology*, Vol. 26, No 1, 2022, pp. 350–363.

Busch, T. and Lewandowski, S., 'Corporate carbon and financial performance: A meta-analysis', *Journal of Industrial Ecology*, Vol. 22, No 4, 2018, pp. 745–759.

CDP, 'Scoring introduction 2023', Tech. rep., CDP, 2023. URL `https://www.cdp.net/en/guidance/guidance-for-companies`.

Cohen, J., 'Statistical power analysis for the behavioral sciences', Academic press, 2013.

Crippa, M., Guizzardi, D., Banja, M., Solazzo, E., Muntean, M., Schaaf, E., Pagani, F., Monforti-Ferrario, F., Olivier, J., Quadrelli, R., Risquez Martin, A., Taghavi-Moharamli, P., Grassi, G., Rossi, S., Jacome Felix Oom, D., Branco, A., San-Miguel-Ayanz, J. and Vignati, E., 'CO2 emissions of all world countries', *Luxembourg: Publications Office of the European Union*, 2022.

Datt, R., Luo, L. and Tang, Q., 'Corporate choice of providers of voluntary carbon assurance', *International Journal of Auditing*, Vol. 24, No 1, 2020, pp. 145–162.

Datt, R., Luo, L., Tang, Q. and Mallik, G., 'An international study of determinants of voluntary carbon assurance', *Journal of International Accounting Research*, Vol. 17, No 3, 2018, pp. 1–20.

Datt, R. R., Luo, L. and Tang, Q., 'The impact of legitimacy threat on the choice of external carbon assurance: Evidence from the us', *Accounting Research Journal*, 2019.

Dutta, P. and Dutta, A., 'Impact of external assurance on corporate climate change disclosures: new evidence from Finland', *Journal of Applied Accounting Research*, Vol. 22, No 2, 2021, pp. 252–285.

ECB/ESRB, 'The macroprudential challenge of climate change', Tech. rep., European Central Bank & European Systemic Risk Board, 2022.

Ekasingh, E., Simnett, R. and Green, W. J., ‘The effect of diversity and the mediating role of elaboration on multidisciplinary greenhouse gas assurance team effectiveness’, *Behavioral Research in Accounting*, Vol. 31, No 1, 2019, pp. 81–96.

Fan, H., Tang, Q. and Pan, L., ‘An international study of carbon information asymmetry and independent carbon assurance’, *The British Accounting Review*, Vol. 53, No 1, 2021, p. 100971.

Galama, J. T. and Scholtens, B., ‘A meta-analysis of the relationship between companies’ greenhouse gas emissions and financial performance’, *Environmental Research Letters*, Vol. 16, No 4, 2021, p. 043006.

Green, W. and Li, Q., ‘Evidence of an expectation gap for greenhouse gas emissions assurance’, *Accounting, Auditing & Accountability Journal*, Vol. 25, No 1, 2011, pp. 146–173.

Green, W. and Taylor, S., ‘Factors that influence perceptions of greenhouse gas assurance provider quality’, *International Journal of Auditing*, Vol. 17, No 3, 2013, pp. 288–307.

Green, W., Taylor, S. and Wu, J., ‘Determinants of greenhouse gas assurance provider choice’, *Meditari Accountancy Research*, Vol. 25, No 1, 2017, pp. 114–135.

Green, W. and Zhou, S., ‘An international examination of assurance practices on carbon emissions disclosures’, *Australian Accounting Review*, Vol. 23, No 1, 2013, pp. 54–66.

GRI, CSR Europe and Accountancy Europe, ‘Member State implementation of Directive 2014/95/EU. A comprehensive overview of how Member States are implementing the EU Directive on non-financial and diversity information’, *Policy & Reporting*, 2017. URL `https://www.accountancyeurope.eu/wp-content/uploads/NFR-Publication-3-May-revision.pdf`.

Griffin, P. A., Lont, D. H. and Sun, E. Y., ‘The relevance to investors of greenhouse gas emission disclosures’, *Contemporary Accounting Research*, Vol. 34, No 2, 2017, pp. 1265–1297.

Hahn, R., Reimsbach, D. and Schiemann, F., ‘Organizations, climate change, and transparency: Reviewing the literature on carbon disclosure’, *Organization & Environment*, Vol. 28, No 1, 2015, pp. 80–102.

Hainmueller, J., ‘Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies’, *Political analysis*, Vol. 20, No 1, 2012, pp. 25–46.

Hair, J. F., Black, B., Babin, B. and Anderson, R., ‘Multivariate data analysis: Global edition (7th editio)’, *Harlow: Pearson Education*, 2010.

Haque, F., ‘The effects of board characteristics and sustainable compensation policy on carbon performance of uk firms’, *The British Accounting Review*, Vol. 49, No 3, 2017, pp. 347–364.

Hassan, O. A. and Romilly, P., ‘Relations between corporate economic performance, environmental disclosure and greenhouse gas emissions: New insights’, *Business strategy and the environment*, Vol. 27, No 7, 2018, pp. 893–909.

He, R., Luo, L., Shamsuddin, A. and Tang, Q., ‘Corporate carbon accounting: a literature review of carbon accounting research from the Kyoto Protocol to the Paris Agreement’, *Accounting & Finance*, Vol. 62, No 1, 2022, pp. 261–298.

IAASB, ‘Handbook of international quality control, auditing, review, other assurance, and related services pronouncements’, Vol. II. International Federation of Accountants, 2021. ISBN 978-1-60815-507-1.

IEA, ‘CO2 emissions in 2022’, Tech. rep., International Energy Agency, 2022.

Imai, K. and Ratkovic, M., ‘Covariate balancing propensity score’, *Journal of the Royal Statistical Society Series B: Statistical Methodology*, Vol. 76, No 1, 2014, pp. 243–263.

ISO, ‘ISO 14064-3: Specification with guidance for the verification and validation of greenhouse gas statements’, International Organization for Standardization (ISO), 2019. ISBN 978-9-26711-193-3. URL `https://www.iso.org/standard/66455.html`.

Jann, B., ‘KMATCH: Stata module module for multivariate-distance and propensity-score matching, including entropy balancing, inverse probability weighting, (coarsened) exact matching, and regression adjustment’. 2017. URL `https://ideas.repec.org/c/boc/bocode/s458346.html`.

Kim, S., Green, W. J. and Johnstone, K. M., ‘Biased evidence processing by multidisciplinary greenhouse gas assurance teams’, *AUDITING: A Journal of Practice & Theory*, Vol. 35, No 3, 2016, pp. 119–139.

KPMG, 'KPMG international: Big shifts, small steps. Survey of responsibility reporting', 2022. URL https://assets.kpmg.com/content/dam/kpmg/se/pdf/komm/2022/Global-Survey-of-Sustainability-Reporting-2022.pdf.

Linden, A. and Samuels, S. J., 'Using balance statistics to determine the optimal number of controls in matching studies', *Journal of evaluation in clinical practice*, Vol. 19, No 5, 2013, pp. 968–975.

Luo, L., Tang, Q., Fan, H. and Ayers, J., 'Corporate carbon assurance and the quality of carbon disclosure', *Accounting & Finance*, 2023.

Michelon, G., Patten, D. M. and Romi, A. M., 'Creating legitimacy for sustainability assurance practices: Evidence from sustainability restatements', *European Accounting Review*, Vol. 28, No 2, 2019, pp. 395–422.

Papadopoulos, G., 'Discrepancies in corporate GHG emissions data and their impact on firm performance assessment', *JRC Working Papers in Economics and Finance*, 2022.

Reuters, 'Volkswagen says diesel scandal has cost it 31.3 billion euros'. 2020. URL https://www.reuters.com/article/us-volkswagen-results-diesel-idUSKBN2141JB. Accessed: March 2023.

Rubin, D. B., 'Using propensity scores to help design observational studies: application to the tobacco litigation', *Health Services and Outcomes Research Methodology*, Vol. 2, 2001, pp. 169–188.

Rubin, D. B., 'The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials', *Statistics in medicine*, Vol. 26, No 1, 2007, pp. 20–36.

Simnett, R., 'Assurance of sustainability reports: Revision of ISAE 3000 and associated research opportunities', *Sustainability Accounting, Management and Policy Journal*, 2012.

Stuart, E. A., Lee, B. K. and Leacy, F. P., 'Prognostic score–based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research', *Journal of clinical epidemiology*, Vol. 66, No 8, 2013, pp. S84–S90.

UNEP, 'Emissions Gap Report (EGR) 2022: The Closing Window – Climate crisis calls for rapid transformation of societies', 2022.

Velte, P., Stawinoga, M. and Lueg, R., 'Carbon performance and disclosure: A systematic review of governance-related determinants and financial consequences', *Journal of Cleaner Production*, Vol. 254, 2020, p. 120063.

Venter, E. R. and Van Eck, L., 'Research on extended external reporting assurance: Trends, themes, and opportunities', *Journal of International Financial Management & Accounting*, Vol. 32, No 1, 2021, pp. 63–103.

Wooldridge, J. M., 'Econometric analysis of cross section and panel data', MIT press, 2010.

WRI and WBCSD, 'The GHG Protocol: A corporate accounting and reporting standard', *World Resources Institute and World Business Council for Sustainable Development*, 2004. URL https://ghgprotocol.org/corporate-standard.

Zhang, Z., Kim, H. J., Lonjon, G. and Zhu, Y., 'Balance diagnostics after propensity score matching', *Annals of translational medicine*, Vol. 7, No 1, 2019.

Zhou, S., 'Reporting and assurance of climate-related and other sustainability information: A review of research and practice', *Australian Accounting Review*, Vol. 32, No 3, 2022, pp. 315–333.

Zhou, S., Simnett, R. and Green, W. J., 'Assuring a new market: The interplay between country-level and company-level factors on the demand for greenhouse gas (GHG) information assurance and the choice of assurance provider', *Auditing: A Journal of Practice & Theory*, Vol. 35, No 3, 2016, pp. 141–168.

List of figures

List of tables

Annex: Literature, distributional aspects, sectoral patterns & robustness checks

Annex 1. Literature evidence

Table 2 presents a description of the independent variables included in the analysis along with studies which document the existence of a statistically significant association between them and the dependent variables of interest.

Table 2: Literature on statistically significant correlates of company emissions and carbon disclosure assurance.

Variable	Short description	Studies documenting statistically significant link	
		with corporate emissions	with external assurance
LOGA	Logarithm of total assets, as a proxy of company size	(Griffin et al., 2017) (Haque, 2017) (Hassan and Romilly, 2018) (Alam et al., 2019) (Bolton and Kacperczyk, 2021a) (Bolton and Kacperczyk, 2021b)	(Zhou et al., 2016) (Datt et al., 2018) (Datt et al., 2019) (Fan et al., 2021)
ROA	Ratio of yearly revenues to total assets, as a proxy of company profitability	(Alam et al., 2019) (Bolton and Kacperczyk, 2021a) (Bolton and Kacperczyk, 2021b)	(Zhou et al., 2016) (Datt et al., 2018) (Fan et al., 2021)
CAPXA	Ratio of capital expenditures to total assets, as a proxy of company capital intensity	(Haque, 2017) (Bolton and Kacperczyk, 2021a) (Bolton and Kacperczyk, 2021b)	(Fan et al., 2021)
LOGPPE	Logarithm of gross plant, & property equipment as a proxy of company physical capital	(Haque, 2017) (Bolton and Kacperczyk, 2021a) (Bolton and Kacperczyk, 2021b)	(Zhou et al., 2016)
LEV	Ratio of total debt to total assets as a proxy of company leverage	(Alam et al., 2019) (Bolton and Kacperczyk, 2021a) (Bolton and Kacperczyk, 2021b)	(Datt et al., 2018)
EMINS ([1])	Dummy if a company belongs to an emissions intensive sector (Utilities, Materials, Energy)	(Hassan and Romilly, 2018)	(Fan et al., 2021)

Note: Natural logarithm applied in relevant transformations.
([1]) In addition to the two studies mentioned in the Table, the rest include sector fixed effects, except (Haque, 2017).

As seen in Table 2, many variables are associated to both corporate emissions and carbon disclosure assurance. Moreover, several studies confirm the same, statistically significant, links thus supporting the robustness of the identified relationships. It should be noted that, although not mentioned in Table 2, virtually every study cited includes static, country-specific factors with statistically strong connection to the respective dependent variable.

Table 3 below presents the descriptive statistics of the control variables used in the analyses.

Table 3: Descriptive statistics of control variables.

Statistic	$LOGA$	ROA	$CAPXA$	LEV	$LOGPPE$	$EMINS$
Minimum	16.20	-0.09	0.00	0.00	10.90	0.00
Maximum	28.60	6.71	0.48	2.56	26.69	1.00
Mean	22.48	0.67	0.04	0.27	20.83	0.18
Median	22.34	0.58	0.03	0.25	20.91	0.00
Std. Dev.	1.90	0.58	0.04	0.19	2.24	0.38
Observations	5240	5240	5001	5240	4774	5254

Note: Natural logarithm applied in relevant transformations.
Source: Refinitiv Eikon and author's calculations.

An interesting observation is that means are close to medians for every continuous variable. Also, variation is more pronounced in profitability (*ROA*), capital expenditure (*CAPXA*) and leverage (*LEV*) while it is less evident in total (*LOGA*) and fixed assets (*LOGPPE*) Finally, the average of *EMINS* indicates that 18% of the whole sample belongs to the emissions intensive sectors of utilities, materials or energy.

Annex 2. Coverage of emissions assurance

In the original sample there are cases where only a fraction of a company's total emissions has been assured. This could be due to specific regulatory requirements (e.g. only part of a company's emissions might fall under the scope of EU ETS and its mandatory assurance). Thus, the high associated costs might lead some firms to avoid engaging with assurance beyond what is legally required.

This could interfere with the analysis of assurance's effect on carbon disclosures. In particular, the impact of falsely classifying such observations as assured would depend on the direction and magnitude of misestimation in the non-assured part of reported emissions. Therefore, observations of which the proportion of total emissions assured is below 95% have been excluded from the analysis.

Figure 9 and Figure 10 present the distribution of cases with low assurance coverage for each emissions Scope.

Figure 9: Distribution of low-coverage Scope 1 emissions assured.

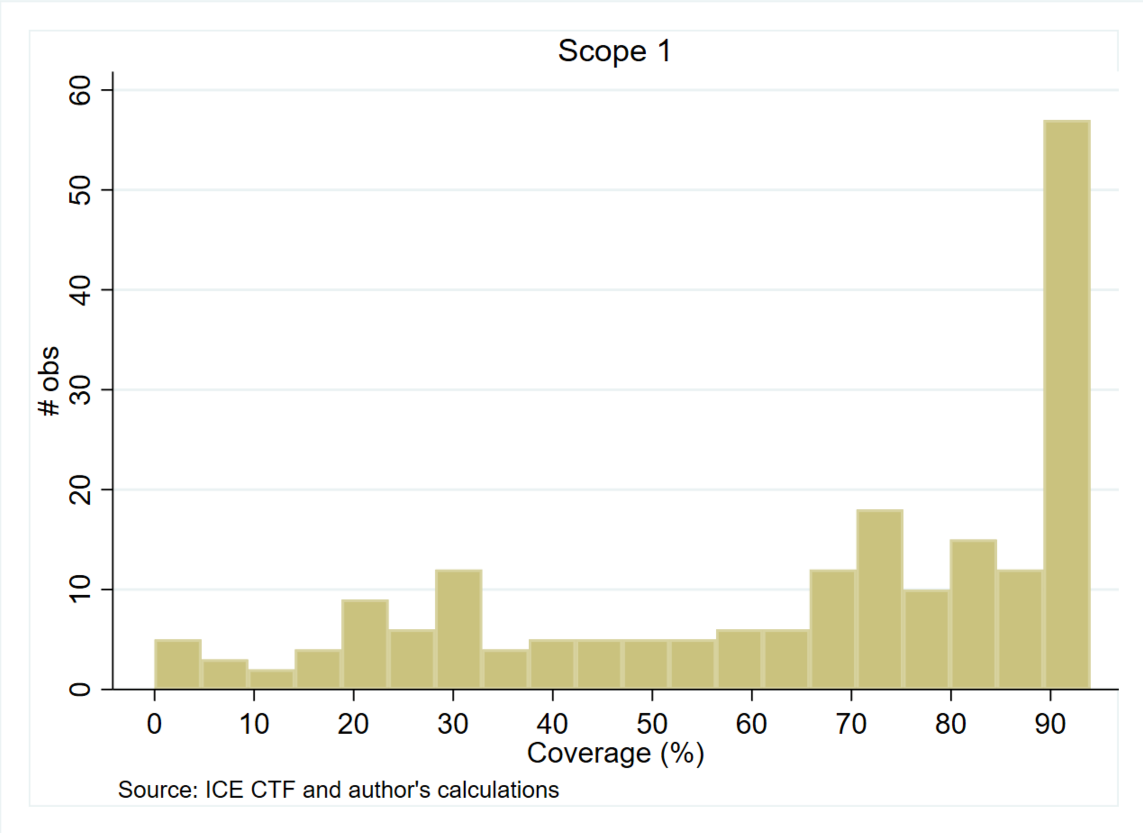
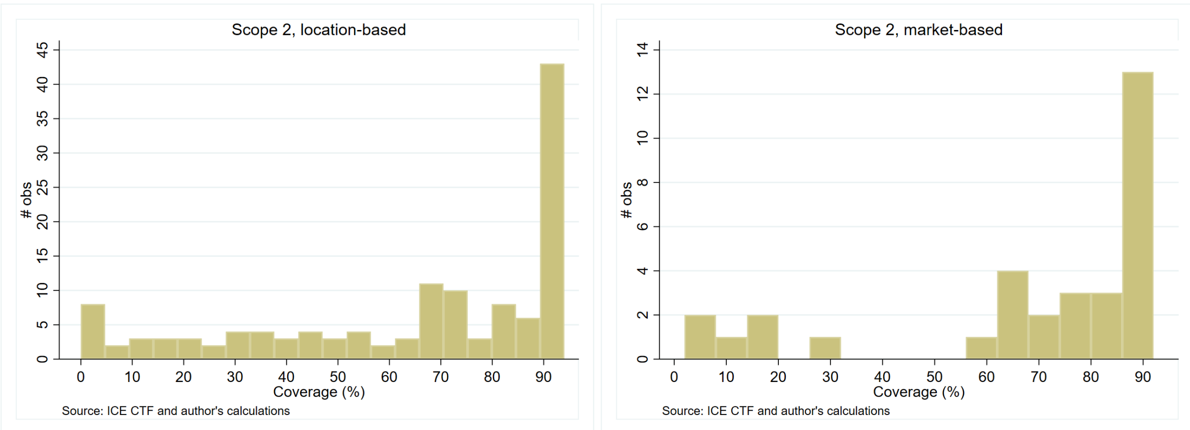


Figure 10: Distributions of low-coverage Scope 2 emissions assured.



The graphs in Figures 9 and 10 show that the low-coverage observations are a very small part of the original sample. Scope 1 emissions contain most of them, almost 4% of the respective sample, while Scope 2 emissions have even fewer cases. About 2.6% of total location-based and even less, 1.1%, of market-based Scope 2 reported emissions data exhibit assurance coverage below 95%.

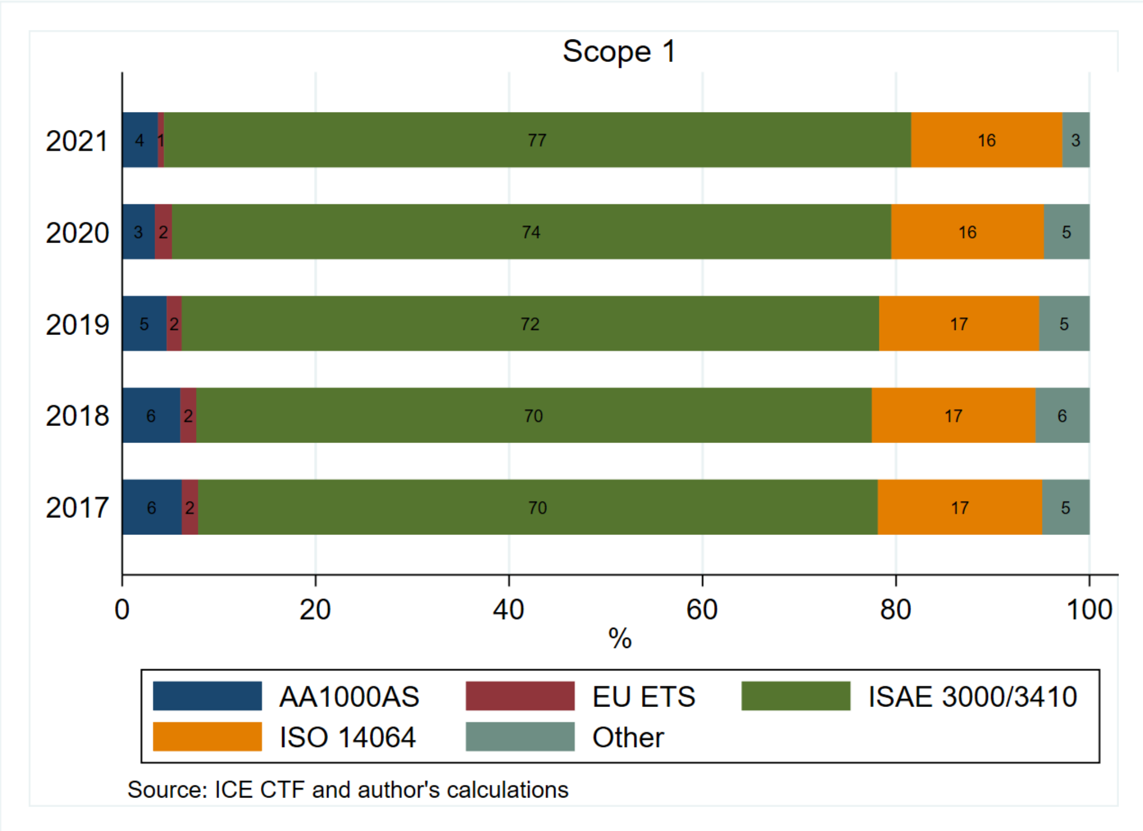
This pattern is possibly related to the nature of each Scope's estimation. Market-based Scope 2 figures need to be accompanied by certain documents (e.g. energy purchase contracts or certificates), hence it is easier to assure them. On the other hand, part of companies' Scope 1 emissions are under the scope of EU ETS. Thus, a company might decide to assure only the legally required share of its global emissions but not the remaining one.

After removing these observations, the sample's average assurance coverage is 99.2% for Scope 1, 99.4% for location-based Scope 2 and 99% for market-based Scope 2 emissions.

Annex 3. Applied carbon assurance standards

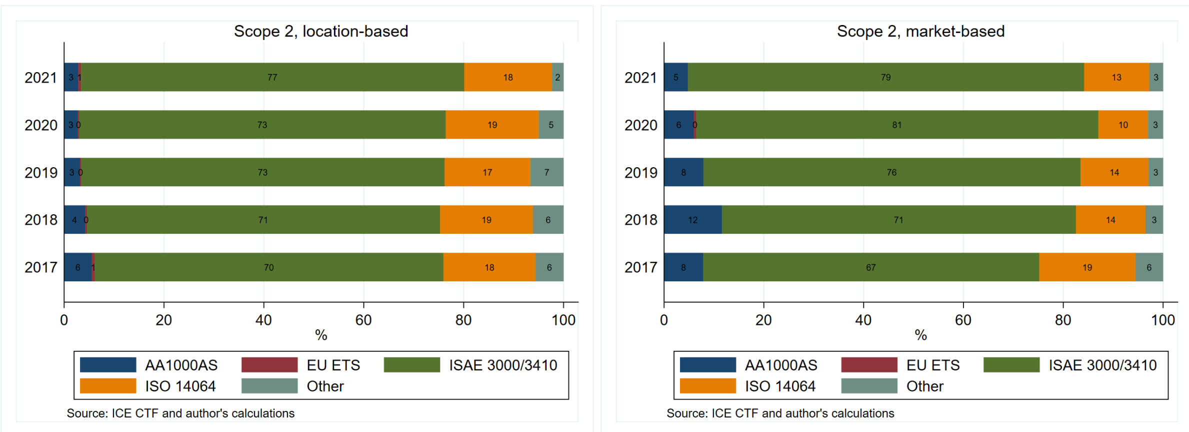
An interesting aspect of carbon assurance engagements is the multitude of applied standards. While it is not the current study's main focus, the data permit an initial examination of the recent trends in this aspect of carbon assurance practice in Europe.

Figure 11: Standards used for Scope 1 emissions' assurance by year.



Frequency of each assurance standard's application as a share of total assured information per year.

Figure 12: Standards used for Scope 2 emissions' assurance by year.



Frequency of each assurance standard's application as a share of total assured information per year.

The graphs in Figures 11 and 12 depict each standard's application share in assurance engagements. They show that the most widely used standard is by far the International Standard on Assurance Engagements (ISAE) 3000 and its dedicated version for GHG emissions disclosures, ISAE 3410, or some variation based on them (e.g. ASAE 3000 or the Dutch Standard 3000A) (IAASB, 2021). The second and third most commonly applied standards are ISO 14064 (ISO, 2019) and AA1000AS (AccountAbility, 2020), respectively. The latter's share declines over time in favor of ISAEs 3000/3410, whose gradually increases. Several studies identified ISAE

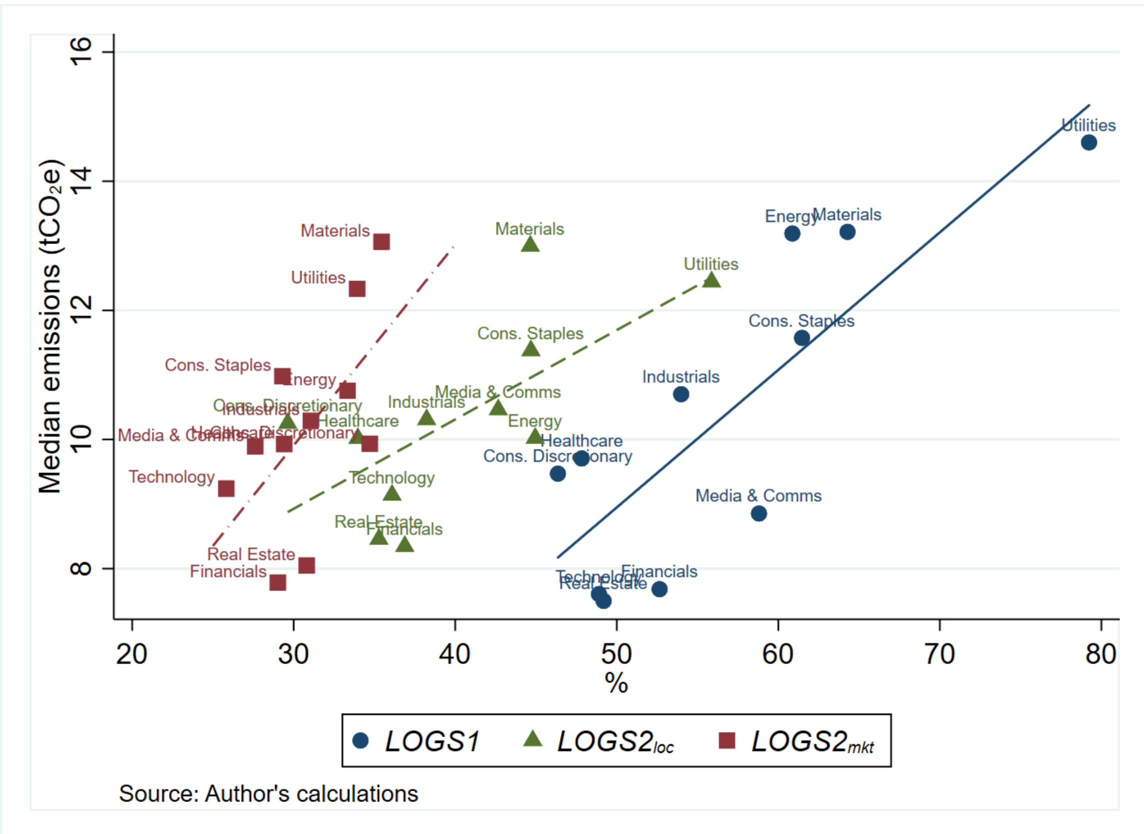
3000 as the preferred standard used by accounting practitioners when providing assurance services (Simnett, 2012, Green and Zhou, 2013, Ackers and Eccles, 2015). The same has been documented in (Zhou, 2022) who reports the results of a recent survey conducted by the KPMG. The available data in the current study do not allow for an analysis of the reasons behind the observed dominance of ISAEs over the other standards. However, if the identified patterns of the previous literature hold, it could be an indication of growing involvement of accounting professionals in the provision of assurance services in Europe compared to specialist assurance providers.

A notable finding is that in several cases the EU ETS is mentioned as the assurance standard applied although it is not one itself. This information should probably be interpreted that the reported emissions have been verified under the EU ETS Directive and the processes elaborated in it ([23]). Interestingly, despite the fact that the EU ETS's focus is on direct emissions, a tiny fraction of firms reports its use also for the assurance of indirect, Scope 2 emissions.

Annex 4. Sectoral emissions and assurance association

The initial description of the data reveals a common pattern of heterogeneity in sectoral emissions and the application of assurance in carbon disclosures. The graph in Figure 13 shows a clear representation of the relationship between carbon disclosure assurance (or its lack thereof) and emissions at sectoral level.

Figure 13: Association between sectoral emissions and assurance engagement.



Sector median emissions versus share of assured emissions reports within each sector.

Figure 13 shows that there is a positive link between sectoral emissions and carbon disclosures' assurance. The more(less) polluting a sector is, the higher(lower) the share of assured reports it has. A combination of mechanisms could be behind this pattern. Obviously, it could be that companies in high-emitting sectors are obliged to engage with assurance by law. If such a requirement applies at sector level without any company-specific thresholds then, naturally, this would affect all companies within it.

Another mechanism that might underlie observed patterns is the possible strategic imitation of environmental practices among companies trying to exhibit a level of sensitivity similar to their peers. This mechanism could manifest in sectors where the law does require only some companies to engage with assurance but its more prominent role should be in sectors where carbon disclosure assurance remains absolutely voluntary. In the latter case it could unfold in both directions; either increasing assurance engagements or limiting their application depending on the behavior of the most influential entities.

([23]) See "Verification standards accepted by CDP" in https://www.cdp.net/en/guidance/verification

Annex 5. Main model's diagnostics

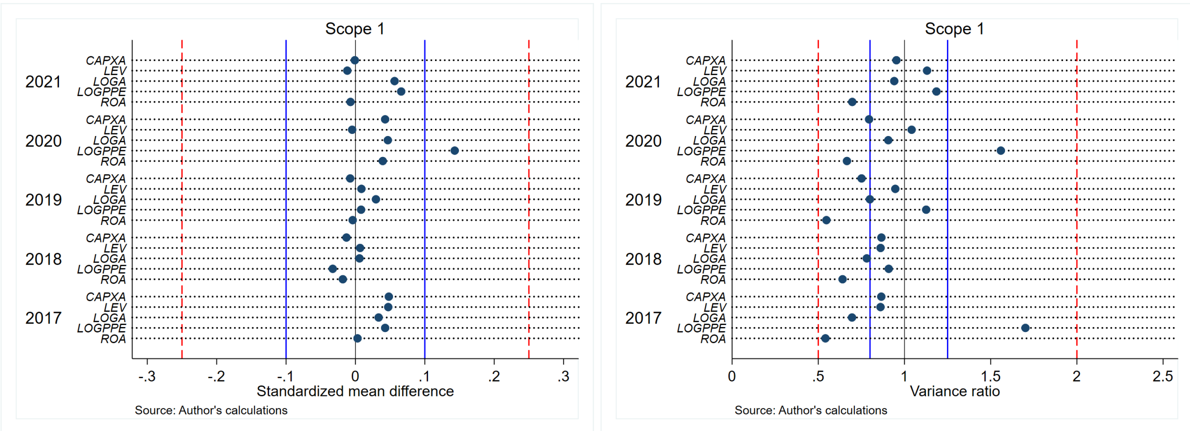
An indispensable element to establish trust in the results is to check the baseline model's diagnostics. These include the assessment of covariate balance between the two treatment groups and the existence of sufficient common support.

To assess whether the treatment and control groups are balanced with respect to covariates, the analysis employs two methods; check standardized mean differences (SMD) and variance ratios (VR) in control variables over treatment groups; and perform a statistical test for covariate balance (Imai and Ratkovic, 2014).

There aren't specific cut-off points to define that balance has been achieved. However, general guidelines in the literature suggest that absolute SMDs below 0.1 or 0.25 are reasonable indicators for balanced covariates and a correctly specified propensity-score model (Rubin, 2001, Austin, 2009, Austin, 2011, Stuart et al., 2013). Some studies point to the similarity between the SMD and Cohen's *d* statistic (Cohen, 2013), proposing that a cut-off of 0.2 can be used to represent a small effect size (Austin, 2009, Linden and Samuels, 2013). With respect to the VR, thresholds recommended by the literature consistent with good balance in covariates are $\frac{4}{5} < VR < \frac{5}{4}$ (Rubin, 2007) or, $\frac{1}{2} < VR < 2$ for acceptably balanced ones (Rubin, 2001, Rubin, 2007, Zhang et al., 2019).

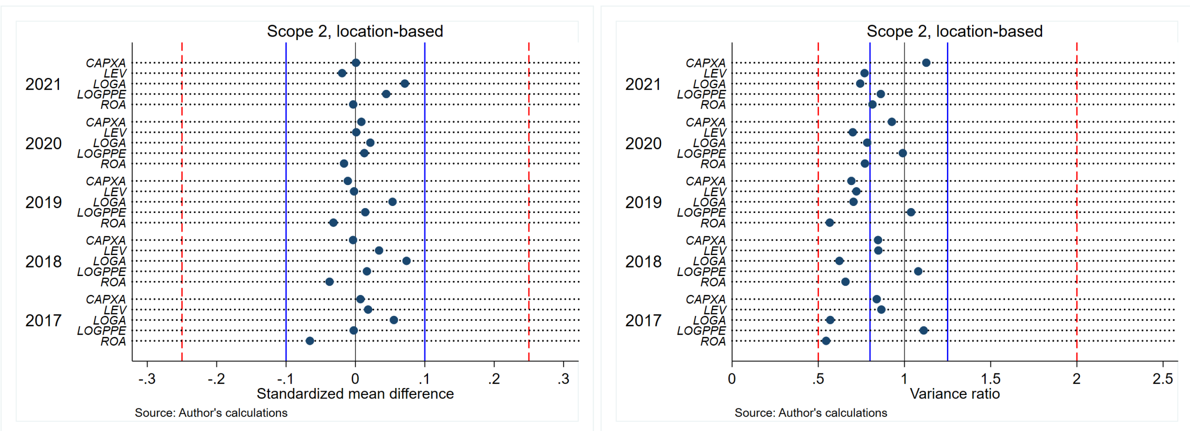
Figures 14 to 16 show the SMDs and VRs in control variables over treatment groups.

Figure 14: Diagnostic statistics of covariate balance in Scope 1 emissions' assurance model, by year.



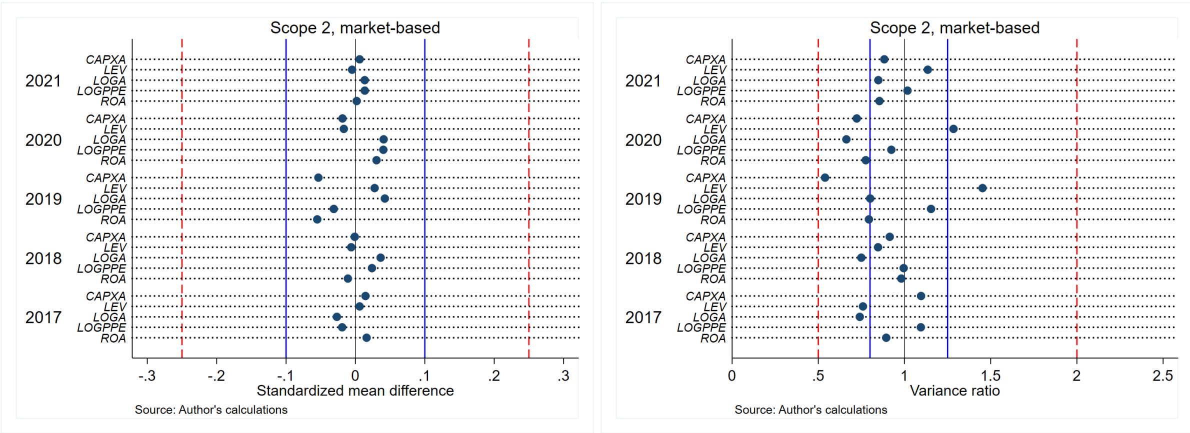
The dashed, red vertical lines and the solid, blue ones denote the range of acceptably and well-balanced covariates, respectively.

Figure 15: Diagnostic statistics of covariate balance in Scope 2, location-based emissions' assurance model, by year.



The dashed, red vertical lines and the solid, blue ones denote the range of acceptably and well-balanced covariates, respectively.

Figure 16: Diagnostic statistics of covariate balance in Scope 2, market-based emissions' assurance model, by year.



The dashed, red vertical lines and the solid, blue ones denote the range of acceptably and well-balanced covariates, respectively.

As shown in Figures 14 to 16, SMDs are, practically, always below the more conservative cut-off points and well within the range of acceptable balance. By comparison, VRs are more dispersed but they also consistently remain in acceptable-balance territory.

In addition to the graphical assessment of covariate balance, the study performs the test described in (Imai and Ratkovic, 2014) to examine whether the propensity score is correctly specified. Table 4 reports the results.

Table 4: Covariate balance test results.

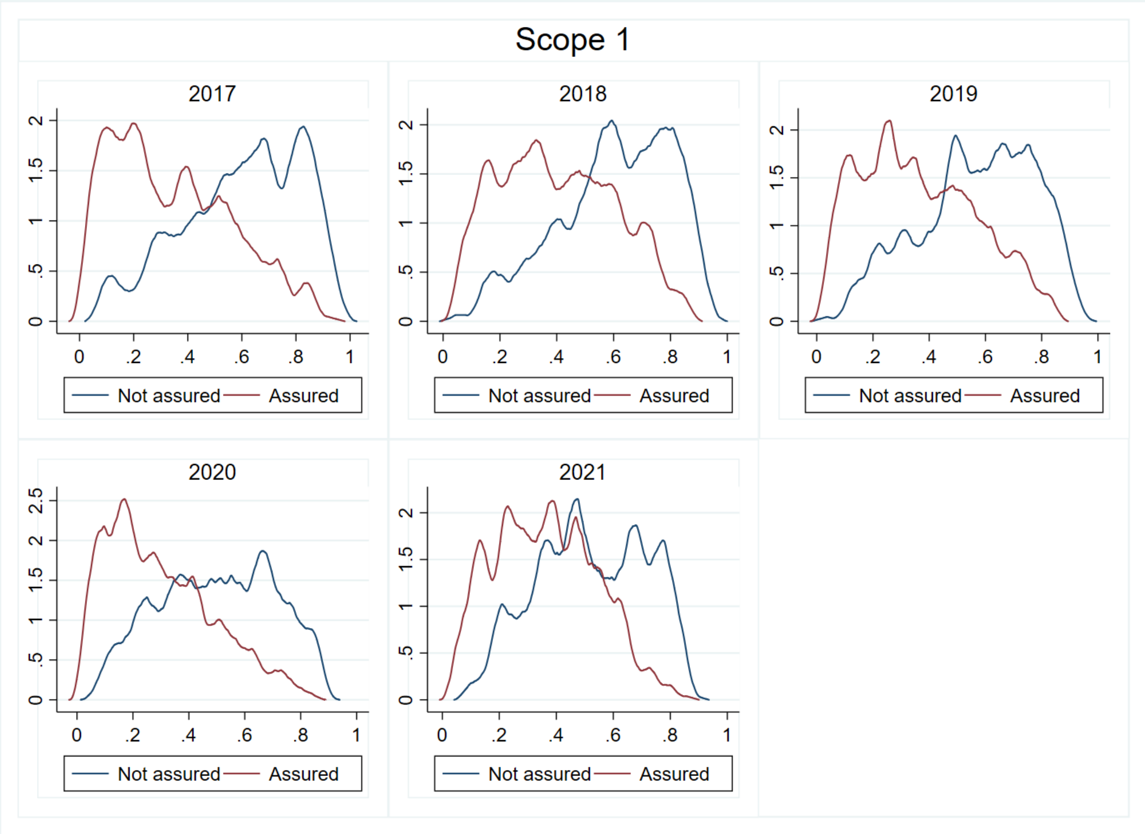
Year	$LOGS1$	$LOGS2_{loc}$	$LOGS2_{mkt}$
2017	0.81	0.13	0.81
2018	0.40	0.31	0.97
2019	0.84	0.06	0.77
2020	0.06	0.71	-
2021	0.18	0.18	0.87

Note: The test's null hypothesis is that covariates are balanced. Reported numbers are the respective p-values. Missing values are due to lack of convergence in the generalized method of moments estimator used to compute the test statistic.

The results in Table 4 corroborate the findings of the previous diagnostic statistics. The estimated p-values do not reject the null hypothesis of covariate balance at standard significance levels. It should be noted that the missing value in Table 4 is due to lack of convergence in the optimization algorithm caused by the presence of some countries with very few observations between treated and control groups.

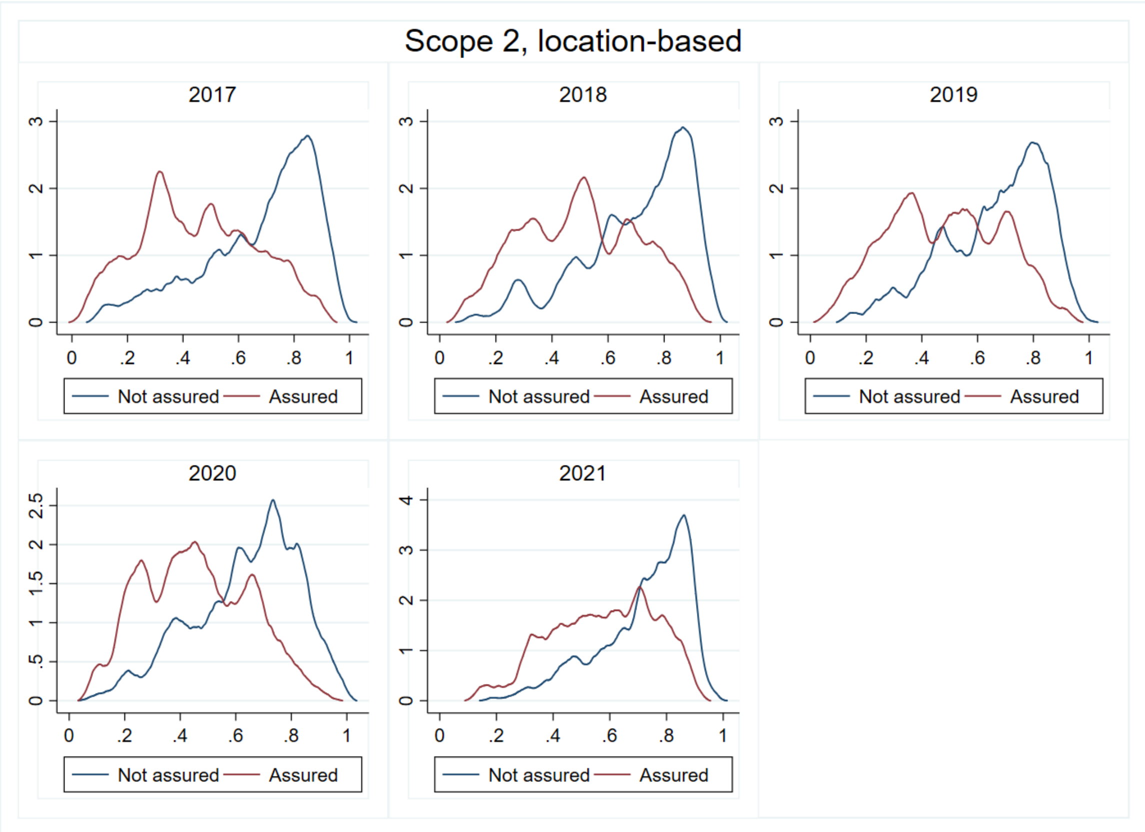
Another important diagnostic is to check whether each individual has a positive probability of receiving each treatment, known as the overlap assumption. Figures 17 to 19 show the densities of the estimated propensity scores at each treatment level for the different models.

Figure 17: Density plots of treatment probability overlap for Scope 1 emissions.



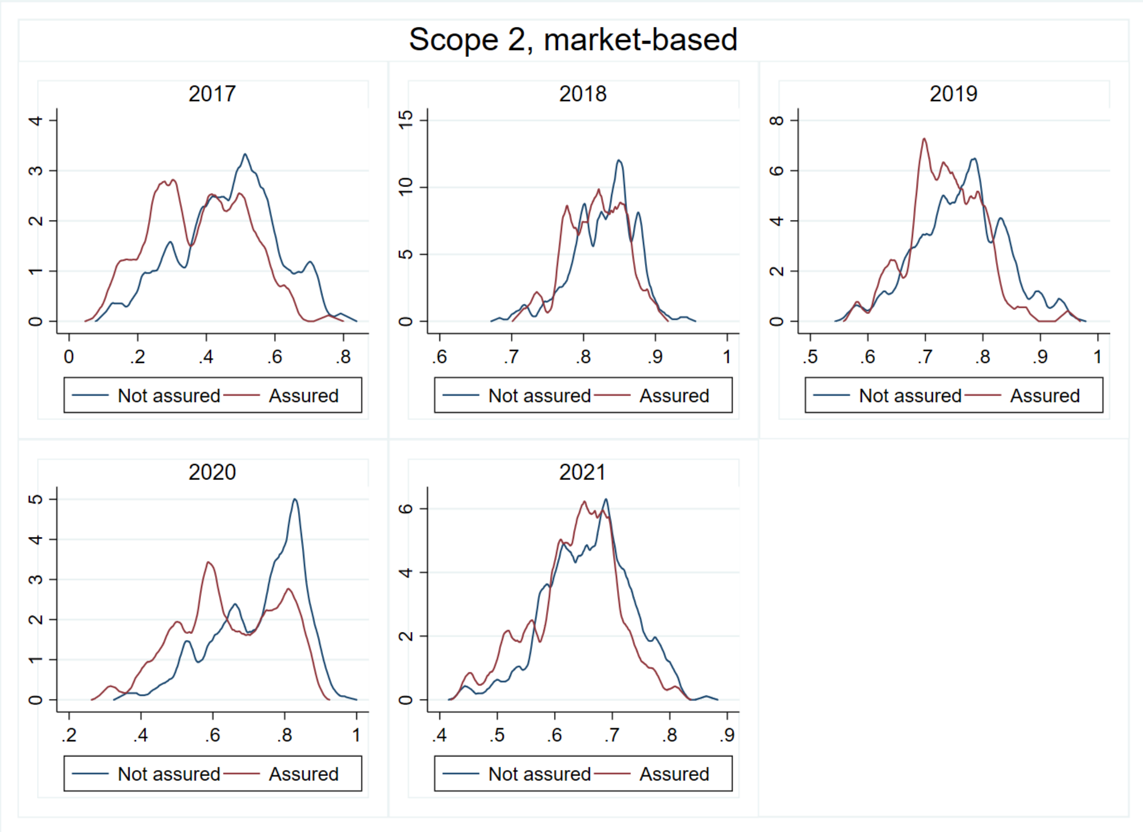
Graphs show densities of estimated propensity scores (horizontal axis), by treatment group.

Figure 18: Density plots of treatment probability overlap for Scope 2, location-based emissions.



Graphs show densities of estimated propensity scores (horizontal axis), by treatment group.

Figure 19: Density plots of treatment probability overlap for Scope 2, market-based emissions.



Graphs show densities of estimated propensity scores (horizontal axis), by treatment group.

The density plots show that, in general, the overlap assumption between the estimated treatment probabilities holds. In most years propensity scores exhibit adequate mass in the region where they overlap without having substantial mass near 0 or 1.

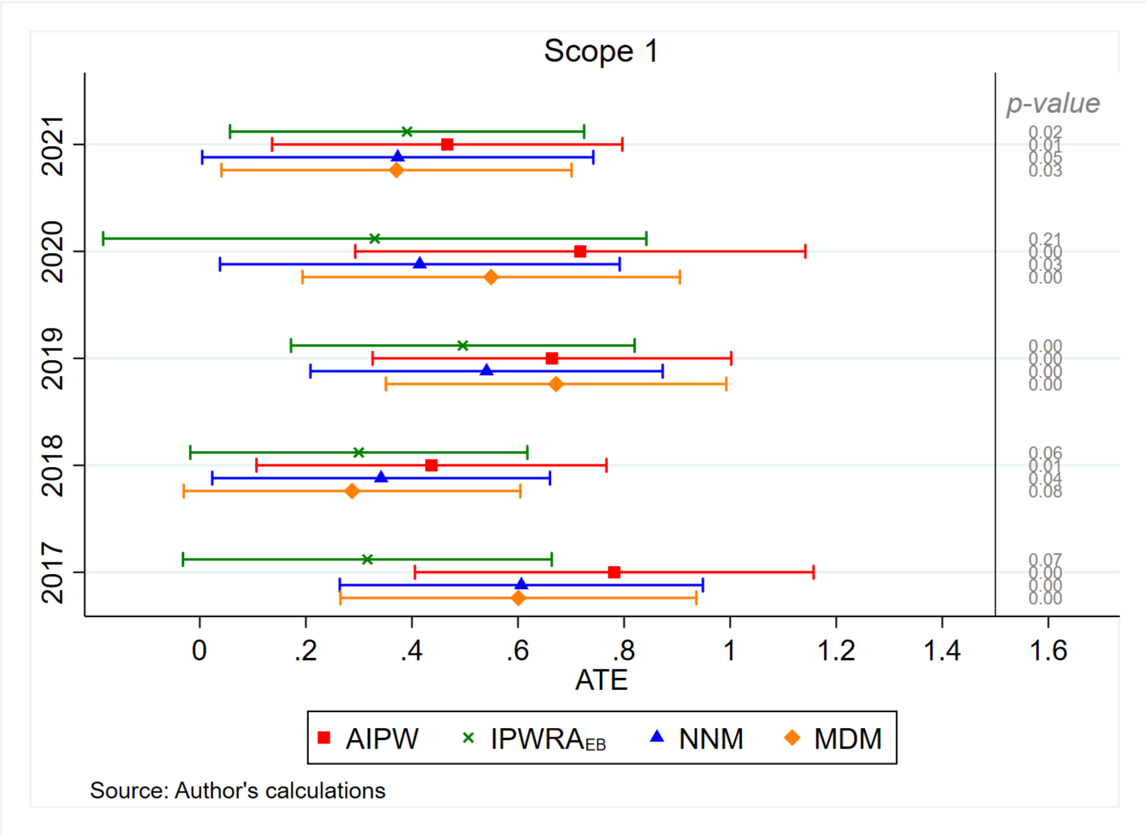
Annex 6. Alternative estimation methods

To guard against possible issues arising from the use of a single estimator or specific functional form dependence, the study examines the results' robustness using 3+1 additional models to estimate ATEs.

These include the augmented inverse-probability-weighted (AIPW) estimator, which is also doubly robust, non-parametric approaches such as nearest-neighbor (NNM) and multivariate-distance kernel matching (MDM) estimators, and the IPWRA estimator with entropy balancing (Hainmueller, 2012). The latter is not a different estimator but rather refines the matching weights by applying entropy balancing to achieve perfect balance in covariates' first and second moments over treatment groups.

In addition to the parametric estimators, such as IPWRA and AIPW, the analysis employs two matching estimators which do not require a specific functional form neither for the treatment nor for the outcome model. The first one is the NNM estimator with replacement and exact matching on the sector-specific variable (i.e. *EMINS*). Four matches per observation (controls to treated subjects) have been used, a number that yields the lowest bias in treatment effect estimates according to the literature (Linden and Samuels, 2013). The MDM estimator applies Mahalanobis-distance matching with an Epanechnikov kernel function to assign larger weights to controls with smaller distances (Jann, 2017). Similarly to the NNM, exact matching by sector group is required. Every estimator in the study uses standard errors clustered at firm level. Figures 20 to 22 present the results from the robustness checks.

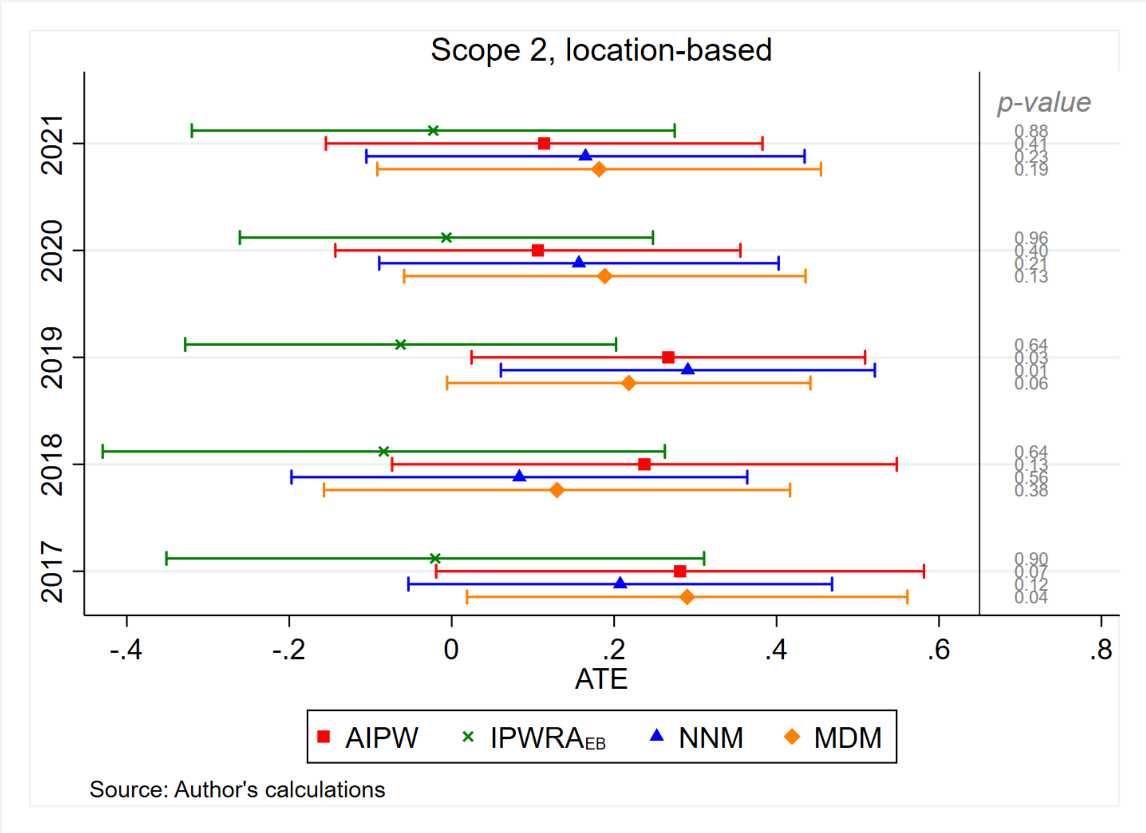
Figure 20: Average treatment effect of external assurance on reported Scope 1 emissions, estimated by various methods.



Note: markers: ATE, whiskers: 95% confidence interval.

For Scope 1 emissions estimated ATEs from every method are always statistically significant, except for 2020 in which the entropy balanced IPWRA estimator's results are not statistically significant. Considering the confidence intervals, the effect's magnitude is comparable to the one identified by the main model with many cases exhibiting quantitatively close point estimates as well. Even when considering the most conservative estimates, usually obtained from the entropy balanced IPWRA estimator, average reported Scope 1 emissions rise by about 3% to 5% when every company assures their reports compared to complete absence of assurance.

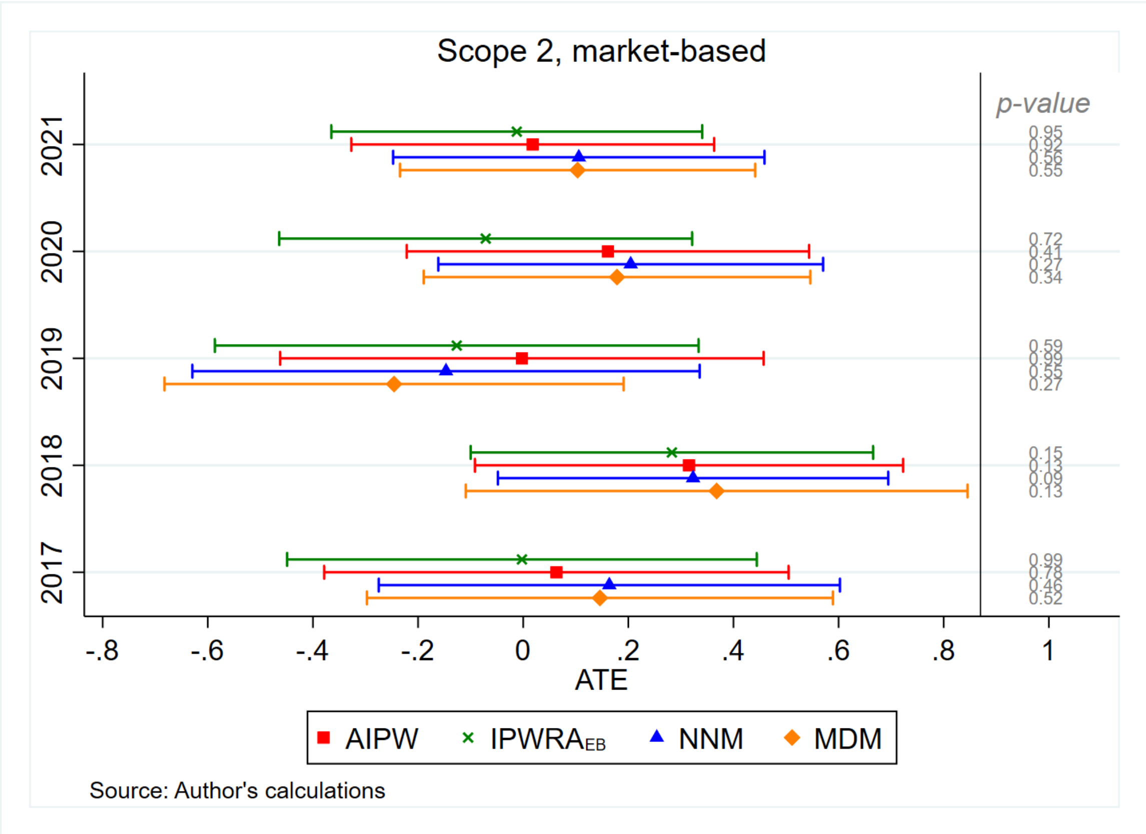
Figure 21: Average treatment effect of external assurance on reported location-based Scope 2 emissions, estimated by various methods.



Note: markers: ATE, whiskers: 95% confidence interval.

Results for location-based Scope 2 emissions reveal that assurance generally does not have a statistically significant effect on reported figures, with the exception of two years. Specifically, in 2017 and 2019, two and three estimators, respectively, yield statistically significant results of similar magnitude. This supports the results from the baseline model pointing to a very subdued effect of assurance on location-based Scope 2 emissions.

Figure 22: Average treatment effect of external assurance on reported market-based Scope 2 emissions, estimated by various methods.



Note: markers: ATE, whiskers: 95% confidence interval.

Finally, robustness check estimates for market-based Scope 2 emissions are completely aligned with the ones from the main model. The only year in which estimators identify a weakly statistically significant effect, is 2018 with the magnitude being roughly the same and close to the baseline model's.

Overall, estimated ATEs from the main and the additional estimators are broadly in line with each other which supports the robustness of the identified effect.

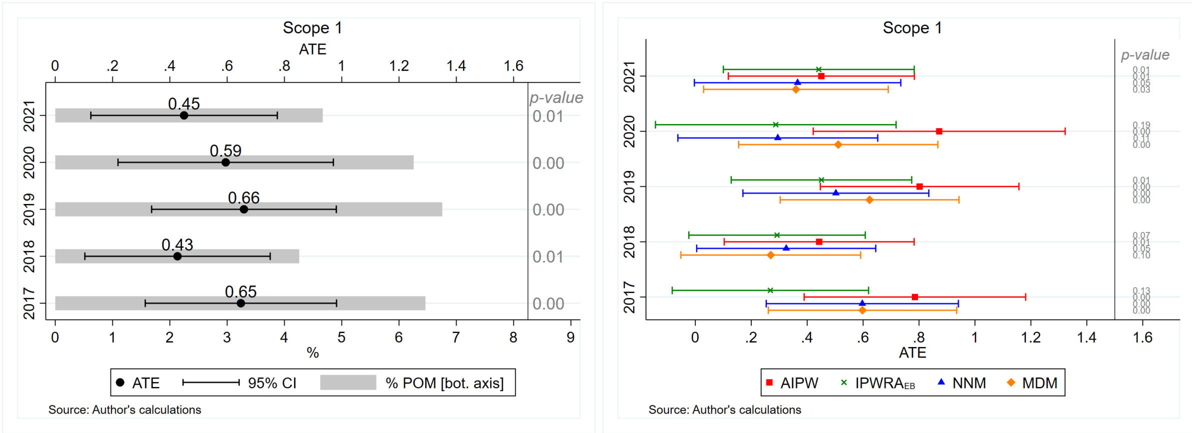
Annex 7. Remove firms reporting EU ETS as verification standard

One specific assumption is that every company has a positive probability of being in either treatment group, i.e. engage with external assurance of its carbon disclosures or not. However, companies under the scope of the EU ETS are obliged to assure their reported emissions hence, by definition, their probability not to assure them would be zero([24]).

To address this concern, companies which have reported EU ETS as the verification standard used (see Figure 11 above), have been removed from the following analysis. This applies only for Scope 1 emissions for two reasons. The main one is that only direct emissions are under the EU ETS's scope, hence verifying indirect, Scope 2 emissions is still voluntary. The second reason is that the number of companies reporting EU ETS as their Scope 2 emissions verification standard is very small (see Figure 12), minimizing the impact of their removal from the sample. Figure 23 shows the results.

([24]) In case they only have facilities under the EU ETS's scope.

Figure 23: Average treatment effect of external assurance on reported Scope 1 emissions, after removing EU ETS firms.



ATEs from the main estimator (IPWRA, left panel) and from alternative ones as robustness check (right panel). Markers always denote ATEs and whiskers their 95% confidence intervals.

The results from the IPWRA and the additional estimators in Figure 23 exhibit the same qualitative, and in many years even quantitative, patterns with the main analysis. The largest observed difference occurs in 2020 in which the estimated ATE is about 1 percentage point lower, in terms of POM share, compared to the baseline estimate. Another interesting observation is that in 2017 and 2020 some of the additional estimators yield statistically insignificant results, slightly above the 10% significance level([25]), similarly to the robustness checks of the main model in Figure 20. Nevertheless, the confidence intervals of the alternative estimators overlap among themselves and with the main model's ATEs.

Annex 8. Endogeneity test

The final robustness analysis checks for potential endogeneity which could arise if unobserved factors that determine reported emissions are correlated with the decision to engage with third-party assurance. This is done using the control-function approach (Wooldridge, 2010) and performing a Wald test to examine whether the estimated correlations between the outcome and treatment models are different from zero. If they are not, it is an indication of endogeneity.

Table 5: Endogeneity test results.

Year	$LOGS1$	$LOGS2_{loc}$	$LOGS2_{mkt}$
2017	0.86	-	-
2018	0.58	-	-
2019	0.90	-	0.50
2020	-	-	0.67
2021	-	-	-

Note: The test's null hypothesis is that treatment and outcome unobservables are uncorrelated. Reported numbers are the respective p-values. Missing values are due to lack of convergence.

The reported p-values in Table 5 suggest that the null hypothesis of no correlation between the unobservables can never be rejected at typical significance levels. Therefore, there is no evidence of endogeneity in the models. It is worth noting that the presence of some countries with very few observations between treated and control groups causes convergence problems to the test's optimization algorithm.

([25]) Only in the case of 2020 and the entropy balanced IPWRA estimator are the results clearly statistically insignificant.

Science for policy

The Joint Research Centre (JRC) provides independent, evidence-based knowledge and science, supporting EU policies to positively impact society