Carbon Accounting Quality:
Measurement and the Role of Assurance

Brandon Gipper
Stanford University Graduate School of Business

Fiona Sequeira
Stanford University Graduate School of Business

Shawn X. Shi
University of Washington Michael G. Foster School of Business

February 2024

Abstract: We examine the role of assurance—third-party verification—on carbon accounting quality. We develop a measure of carbon accounting quality based on the deviation of reported emissions from a model-based expected level and use two other survey-based measures. We show that assurance is associated with improved carbon accounting quality. This association cannot be explained by firm type or firm-level transparency, is isolated to the scope-specific emissions being assured, does not relate to financial reporting quality, and is stronger when assurance is more thorough and pervasive. Assurance improves carbon accounting quality by identifying issues in a firm's carbon accounting system, resulting in fewer omissions and revisions of prior errors. Using the implementation of mandated assurance in three E.U. countries for non-financial reporting, we show that countries with these mandates experience within-firm improvements in carbon accounting quality post-regulation. Together, the findings highlight the importance of external assurance in shaping carbon accounting quality.

JEL Classifications: *G11; G18; G30; M14; M41; M42*

Keywords: *carbon accounting; ESG assurance; sustainability*

Contact: gipperbr@stanford.edu (Gipper), flgs@stanford.edu (Sequeira), shawnshi@uw.edu (Shi). We thank Mallory Bell, Thomas Bourveau, Omer Karaduman, Charles Lee, Maureen McNichols, Sarah McVay, Joe Piotroski, Phil Quinn, Leehi Yona, and workshop participants at Stanford University and the University of Washington for helpful discussion and comments. We thank the Stanford GSB Data, Analytics, and Research Computing group, especially Mason Jiang and Wonhee Lee, for their excellent research assistance. This research is supported by Stanford University's Responsible Investing Initiative. The authors gratefully acknowledge financial support from Stanford University and the University of Washington.

1. Introduction

Combatting climate change is an essential theme around the world, with more than 100 countries—or nearly 50% of world GDP—having carbon net neutrality target commitments (ECIU, 2024). To credibly establish and measure performance against such targets, organizations need to accurately quantify carbon footprints, which gives rise to carbon accounting. In the corporate sector, which released an estimated 49% of emissions in the U.S. in 2021 (EPA, 2023), more firms are disclosing their carbon emissions. There are critical questions regarding the quality of firms' reported carbon emissions and understanding how they can improve their carbon accounting quality. This study develops a measure of carbon accounting quality and examines the role of assurance, i.e., third-party verification, in shaping carbon accounting quality.

We conceptualize carbon accounting quality as the extent to which reported emissions reflect an entity's actual emissions. There are two reasons why a carbon accounting system would not perfectly report an entity's actual emissions. First, the *measurement standard* (e.g., the GHG Protocol) may be flawed (Kaplan and Ramanna, 2021; Glenk, 2023; Reichelstein, 2023). Second, a carbon accounting system inherently involves estimations and judgment, and thus has the potential for errors and biases, i.e., *implementation issues*. For example, a firm may not have the proper metering equipment in all its facilities or the most up-to-date conversion factors to compute its emissions. Because the use of the GHG Protocol is so pervasive across firms, any broad sample, archival assessment of carbon accounting quality can only be performed within this measurement standard. Therefore, we focus our predictions and inferences on implementation issues.

We hypothesize that third-party verification can improve a firm's carbon accounting quality primarily by identifying the firm's implementation issues. The third-party assurance process includes procedures such as checking samples of reported data against source data, interviewing relevant personnel, reviewing internal and external documentary evidence, and conducting on-site visits. Importantly, assurors can engage in "process assurance" which involves reviewing the reporting firm's information systems for collection, aggregation, analysis, and internal verification and review of environmental data. To the extent implementation issues are discovered through assurance and subsequently resolved by reporting firms, we should expect an improvement in carbon accounting quality post-assurance.

However, there are many reasons to believe ESG assurance does not improve carbon accounting quality. First, assurance over emissions metrics is voluntary and lacks regulatory oversight (e.g., lacks the equivalent of PCAOB overseeing financial statement audits) or an assurance standard used by every assuror (Gipper et al., 2023). Second, there exists considerable heterogeneity in the assurance services provided, with more than 90% of the assurance cases being limited assurance, which entails substantially less assuror work than reasonable assurance, the level required for financial statement audits. Anecdotally, some assurors are willing to provide limited assurance with a fee as low as \$15,000—roughly 0.7% of the financial audit fee for an average U.S. public firm in 2021—and simply check spreadsheets shared by the reporting firm. Third, because obtaining assurance allows firms to obtain higher ratings from GRI and CDP, reporting firms may have the incentive to simply "check the box" and do the bare minimum—treating assurance as a mere formality. Because of this tension, whether external assurance improves carbon accounting quality is an empirical question.

The GHG Protocol groups carbon emissions from a firm's operations and activities into three types (or "scopes"), including direct emissions from controlled assets (Scope 1), indirect emissions from energy consumption (Scope 2), and other indirect emissions along the value chain, e.g., by suppliers and customers (Scope 3). We focus on Scopes 1 and 2 emissions because those emissions scopes are within a firm's control to measure and report. We draw inspiration from the earnings quality literature and develop a measure of carbon accounting quality at the firm-year-scope level using an approach analogous to the estimation of abnormal accruals. Specifically, we model reported emissions levels as a function of the firm's economic fundamentals that capture production activities and production technologies; we define "abnormal emissions" as the absolute value of the residual term from the regression. The intuition is that emissions are a byproduct of a firm's operations. Typically, the more a company produces and sells, the more greenhouse gases the firm directly emits (i.e., Scope 1) and the more energy the firm uses and—therefore—the more greenhouse gases it indirectly emits (i.e., Scope 2). If our model is well-specified, deviations from the expected level of emissions are likely driven by implementation issues.

We pair this measure of carbon accounting quality with novel assurance data at the scope level. Our assurance data come primarily from firm disclosures in ESG reports, which are manually collected by Gipper et al. (2023). We supplement this data with assurance information in firms' responses to CDP questionnaires. Using a comprehensive sample of U.S. firms with carbon emissions disclosure from 2010-2020, we document a strong negative association between assurance and abnormal emissions for both Scopes 1 and 2 emissions. We triangulate our inference using two alternate measures of carbon accounting quality from the CDP questionnaires, including management uncertainty in emissions reporting and the time lag between questionnaire release and completion. Both alternative measures negatively correlate with assurance. For all measures, our findings are consistent with the prediction that assurance improves carbon accounting quality.

To mitigate the concern that our finding is driven by correlated omitted variables, we leverage our granular data and show that our finding is unlikely driven by firm type, firms' overall commitment to transparency, or financial reporting quality. We further demonstrate that our results are driven specifically by *carbon* assurance. In "horserace" regressions with assurance indicators for both Scopes 1 and 2 emissions as independent variables, abnormal emissions for Scope 1 strongly associate with assurance over Scope 1 disclosure but do not associate with assurance over Scope 2 disclosure, and vice versa. In addition, we exploit variation in the thoroughness of assurance and show that the result is stronger when a firm obtains assurance for all three scopes of emissions and when the level of assurance is "reasonable"—the same level as a financial statement audit—as opposed to “limited.” Taken together, these results lend credibility to our hypothesis that assurance improves carbon accounting quality.

We explore the mechanism through which assurance improves carbon accounting quality. To do so, we use CDP questionnaire data to develop three specific indicators of the *carbon accounting system* and correlate these measures with assurance. We show that assurance improves carbon accounting quality primarily by helping reporting firms identify issues in their carbon accounting system, which subsequently results in fewer sources of emissions being omitted and historically reported emissions figures being updated.

In our last set of analyses, we generalize our findings to firms with emissions data in 42 countries by documenting a similar association between scope-specific assurance and carbon accounting quality. Further, we use the implementation of the Non-Financial Reporting Directive (NFRD) as a shock to assurance. Adopted by the European Union in 2014 and fully implemented in 2018, NFRD requires certain public companies to provide non-financial disclosure documents. Although NFRD is a disclosure mandate, three member states also opted for mandatory assurance over reporting firms’ sustainability disclosures, creating variation within-E.U. regarding mandated assurance of required emissions disclosures. Using a difference-in-differences design, we show that relative to other E.U. countries, firms in countries with assurance mandates experienced an increase in carbon accounting quality after NFRD implementation.

Our study makes two contributions. First, we contribute to the carbon reporting literature by conceptualizing and developing a new measure of carbon accounting quality (e.g., Downar et al., 2021; Cohen et al., 2023a). Drawing inspiration from the earnings quality literature, we introduce a framework to formally conceptualize carbon accounting quality. We also develop a novel measure (i.e., abnormal emissions) that captures the *quality* of reported emissions figures rather than the amount of environmental disclosure, is independent of coverage by data vendors for vendor-created reporting quality measures, is specific to each scope of emissions, and can be easily estimated as long as emissions disclosure exists.[1] Furthermore, prior studies focus primarily on Scope 3 emissions while assuming that Scopes 1 and 2 emissions are fairly accurate. By introducing our measures and studying their associations with assurance, we highlight reporting quality issues surrounding Scopes 1 and 2. Our measure of carbon accounting quality should be of interest to investors, regulators, and other stakeholders. Carbon emissions levels and intensity are generally used by practitioners to evaluate firm exposure to carbon-related risks, which often leads to divestments or changes in the cost of capital (e.g., Bolton and Kacperczyk, 2021). Understanding the quality of emissions disclosures will likely allow for more informed stakeholder decision-making, such as more efficient allocation of resources and risks.

Second, we contribute to a literature that examines the emergence and voluntary adoption of verification services, including auditing and other, new forms of assurance of non-financial information or processes (e.g., Watts and Zimmerman, 1983; Minnis, 2011; Duflo et al., 2013; Schoenfeld, 2022; Bourveau et al., 2023a, b). Our study is particularly related to ESG assurance, of which the U.S. has seen a striking increase over the last decade. Because ESG assurance is voluntary, lacks regulatory oversight, is primarily performed at a "limited" level as opposed to reasonable, and is highly heterogeneous in many dimensions, it is unclear whether such assurance is useful for firms and stakeholders (Gipper et al., 2023). We focus on assurance over emissions metrics and demonstrate its usefulness in improving carbon accounting quality.[2]

---

[1] Our study also contributes to the literature that examines the quality of ESG reporting more broadly, e.g., disclosure scores from Bloomberg and Refinitiv (Lopez-de-Silanes et al., 2019), restatement of metrics in ESG reports (e.g., Ballou et al., 2018; Michelon et al., 2019), integrated reporting ratings from practitioners such as Ernst & Young (Barth et al., 2017; Maroun, 2019), and disclosure indices developed by researchers (Plumlee et al., 2015). Those measures typically focus on the *amount* of disclosure (e.g., the number of ESG metrics disclosed) and consider both *qualitative* and quantitative disclosures.

[2] Our study is related to papers that study the role of assurance in ESG reporting quality more broadly. Using scores assigned by Ernst & Young, Maroun (2019) shows that external assurance increases the quality of integrated reports—which include both financial and other value-relevant information—in South Africa. Ballou et al. (2018) and Michelon et al. (2019) document a positive association between assurance in ESG reports and the probability of a reporting firm revising its reported numbers—regardless of metric type—in subsequent years. Luo et al. (2023) find a positive association between carbon assurance and CDP disclosure scores. However, assurance has been explicitly incorporated into CDP ratings since 2017, so a positive association between assurance and CDP ratings would likely be mechanical.

The rest of the paper proceeds as follows. The next section introduces a conceptual framework of carbon accounting quality and describes our main hypothesis. Section 3 develops a measure of carbon accounting quality and outlines the empirical strategy. The main empirical results of the paper are provided in section 4, and additional analyses using non-U.S. data are presented in section 5. Section 6 concludes.

2. Conceptual Underpinning

2.1 Reporting Emissions

Emissions data are most often measured and reported to follow the Greenhouse Gas (GHG) Protocol. The GHG Protocol distinguishes between three emission sources labeled Scopes 1-3. Scope 1 emissions are "direct" emissions from sources that are controlled or owned by a firm. Scope 2 and Scope 3 emissions are "indirect"—meaning they are not produced by the company itself. Scope 2 emissions are associated with the purchase of electricity, steam, heat, or cooling. Scope 3 emissions cover the entire value chain from sources not controlled by the company, such as emissions due to the use of the firm's products or employees' work-related travel.[3]

The procedures of the corporate GHG Protocol can be summarized in three steps. The first step is to choose the organizational boundary—either equity share or control—for consolidating GHG emissions. This step is somewhat analogous to the concept of consolidation used in financial reporting in determining what to include or not include in financial statements. The second step is to choose the operational boundaries, i.e., how emissions can be classified into Scopes 1, 2, and 3. The third step is to calculate GHG emissions. The common procedure is to identify all emissions

[3] Some prior studies focus on the reporting quality of Scope 3 emissions while implicitly assuming that Scopes 1 and 2 emissions are accurately measured (e.g., Klaaben and Stoll, 2021). In many industries, Scope 3 emissions comprise most firms' total emissions under the GHG Protocol. Of course, Scope 3 (and Scope 2) emissions are the Scope 1 emissions of other entities, leading to a double counting critique of the GHG Protocol. Moreover, the reporting quality of a firm's Scope 3 emissions depends critically on the reporting quality of the equivalent of Scopes 1 and 2 emissions from entities along its value chain, including non-corporate entities, like employees. For this reason, there is a large amount of estimation and assumptions in reported Scope 3 emissions. Implementation issues for Scope 3 emissions can be orthogonal to implementation issues measuring and reporting Scopes 1 and 2 emissions.

sources within the chosen boundaries, multiply a measure of activity with a corresponding emissions conversion factor for each emissions source, and aggregate the estimated emissions. The common practice is to express emissions in carbon dioxide equivalents (CO_2e), meaning that non-CO_2 greenhouse gas emissions (e.g., methane) are converted to CO_2e using the appropriate conversion factors.[4]

2.2 Conceptualizing Carbon Accounting Quality

We conceptualize carbon accounting quality as the extent to which reported emissions reflect an entity's actual level of emissions.[5] Our conceptualization of carbon accounting quality mirrors the framework proposed by Dechow et al. (2010) to analyze earnings quality. Dechow et al. (2010) argue that reported earnings is a *function* of an enterprise's financial performance during a reporting period. Earnings quality is an evaluation of this function, which is essentially an accounting measurement system that establishes a mapping between reported earnings and financial performance. The authors point out three factors for why reported earnings would not perfectly measure financial performance. The first factor is multiple decision models, i.e., the ability of earnings to summarize a firm's unobservable financial performance for decision-making. An accounting system that produces a single reported earnings number cannot properly represent the underlying financial performance that is relevant for all decision-makers. The second factor is measurement standards. No standard can perfectly measure performance for any given firm. The third factor is the implementation of an accounting system, which inherently involves estimations and judgment.

---

[4] Firms following the GHG Protocol are required to disclose the reporting boundaries, reporting period, Scopes 1 and 2 emissions, and calculation methodologies, among others. Importantly, the GHG Protocol requires firms to disclose emissions independent of any offsets and allowances, which are certificates that allow firms to mitigate the influence of their emissions through forest protection, installation of direct air capture facilities, use of renewable energy sources, etc. Like the organizational boundary and consolidation, offsets in carbon accounting and netting in financial accounting are conceptually similar.

[5] We adopt phrasing where emissions are converted to carbon equivalents in practice, i.e., normalizing any air-based emissions to the prevalent and impactful greenhouse gas carbon dioxide.

These three factors from Dechow et al. (2010) which influence earnings quality can similarly apply to carbon accounting quality. First, in the context of carbon accounting, we believe that actual emissions track a firm's carbon footprint in a straightforward way, but reported emissions produced by the carbon accounting system may not be equally relevant in all decisions. Dechow et al. (2010) note that an accounting system would not perfectly measure financial performance due to many decision models that could utilize the measure. For example, earnings quality could be low for stewardship purposes (e.g., as a bonus plan metric) but high for valuation purposes (e.g., as an input to an investor's enterprise value calculation). Similarly, a carbon accounting system could generate a measure that would be an input to multiple decisions, such as meeting a "net zero" commitment or quantifying the exposure to regulatory risk arising from cap and trade or a governmental entity's emission reduction policies. We do not specifically address the multiple decision model factor for carbon accounting quality but believe this to be a potentially interesting area for future research.

Second, the *measurement standards* (e.g., the GHG Protocol) could be flawed. For example, a widely discussed case is the GHG Protocol's Scope 3 emissions, which are emissions measured along the value chain. Research in measurement standards focuses predominately on Scope 3 emissions. For example, Kaplan and Ramanna (2021) contend that "the same emissions are reported multiple times by different companies, while some entities entirely ignore emissions from their supply and distribution chains." In addition, Glenk (2023) argues that the GHG Protocol fails to represent a firm's emissions faithfully. However, the use of this standard—GHG Protocol—is pervasively used across firms and is not updated frequently, so there is almost no variation to understand its impact empirically, leading to primarily theoretical critiques. Thus, we do not address the impact of measurement standards on carbon accounting quality, but we believe that our approach can be applied to standard changes or similar analyses that may arise in the future.

Third, a carbon accounting system inherently involves estimations and judgment, and thus has the potential for errors and biases, i.e., *implementation issues*. CDP classifies implementation issues into several non-mutually exclusive groupings. We use these groups to illustrate those issues. We provide anecdotes pertaining to each group in Appendix B.

Data Management. Unlike financial accounting, where firms use established accounting software to track transactions and produce financial statements, the use of carbon accounting software is less common. Many firms track and calculate emissions using Excel spreadsheets. Even within a firm, establishments record data using different Excel templates with different data granularity. As a result, errors could arise from data entry, aggregation, and incorrect Excel formulas, among others. Data management issues could also arise when firms choose emissions conversion factors, such as using incorrect or outdated factors.

Data Gaps. Some firms do not have the proper information system in place to gather emissions data from certain sources. For example, a firm may not receive natural gas bills for leased facilities, and if the firm does not actively track down such information from the landlords, data gaps would exist. As another example, multinational firms may lack emissions data from certain overseas offices and establishments. Instead of estimating by making assumptions, some firms simply omit such offices and establishments from their reporting.

Assumptions and Extrapolation. When data gaps exist, firms may either ignore data gaps (which leads to underreporting) or make assumptions and extrapolate. For example, when exact metered data for leased space is not available, a firm could extract a pro-rata share. However, some landlords do not have complete data, have inconsistent data, or are unresponsive to firm requests. In those cases, firms need to make further assumptions, such as using kWh electricity or natural gas per square foot averages in the region to extrapolate. In some cases, firms simply use the number from the previous reporting period.

Sampling. Some firms do not measure Scopes 1 and 2 emissions for all sources; instead, they obtain some samples and extrapolate. Similarly, sampling is a common technique to generate data for Scope 3 emissions. Because the quantity of data that needs to be collected is large, a firm may find it impossible to collect data from each activity in a Scope 3 category. In those cases, a firm may use sampling to extrapolate data from a representative sample of activities within the category. For example, if a company wants to collect data on employee commuting (Category 7 of Scope 3 emissions), instead of collecting data for every single employee, the company may sample a subset of employees and extrapolate. Improper sampling techniques could introduce significant measurement errors.

The above-discussed cases largely abstract away from managerial incentives to manipulate emissions. This contrasts with the earnings quality literature, in which implementation issues include both innocuous errors and manipulations by managers. While our discussion above focuses primarily on errors, manipulation could theoretically exist. Such incentives could stem from CEO compensation being tied to emissions reduction goals. For example, when multiple conversion factors are available, a CEO could instruct her subordinates to choose the factor that results in the lowest calculated emissions. Another example, because emissions are aggregated from all sources within the reporting boundary and firms have the discretion to define their reporting boundary, a firm could strategically choose a reporting boundary that results in lower reported emissions.

2.3 Assurance and Carbon Accounting Quality

By assurance, we refer to third-party verification of firm-disclosed ESG metrics, specifically GHG emissions. Assurance over emissions resembles a financial statement audit but differs in many dimensions. We refer interested readers to Gipper et al. (2023) for a more detailed discussion. In general, ESG assurance is voluntary in most countries, including the U.S. The assurance providers include both traditional auditors who also perform financial statement audits and ESG engineering and consulting firms with specialties in environmental and social matters. In the U.S., roughly 80% of assurance services are provided by non-financial assurors. In terms of assurance quality, the contracted level of assurance is mostly "limited assurance," which is somewhat analogous to a "review" engagement for quarterly financial reports in the U.S. Reasonable assurance, which is the level of assurance required for annual financial statement audits, is very uncommon in the non-financial space. Assurance providers use a variety of assurance standards, including AICPA, ISAE, and ISO, among others. In the aggregate, although assurance over ESG metrics is increasingly common, there exists considerable heterogeneity in many dimensions compared to financial audits.

Our conceptual framework of carbon accounting quality entails both measurement standards and implementation issues. However, the measurement standard—GHG Protocol—is pervasively used; thus, as we argue above, an archival assessment of carbon accounting quality can only be done within this measurement standard. Therefore, we hypothesize that assurance can increase a firm's carbon accounting quality by identifying *implementation issues* in reporting firms' carbon accounting systems. Although there is significant heterogeneity in what an assurance engagement entails, the third-party assurance process could include procedures such as checking samples of reported data against source data, interviewing relevant personnel, reviewing internal and external documentary evidence, and conducting on-site visits. Assurors also engage in "process assurance" which involves reviewing the reporting firm's information systems for collection, aggregation, analysis, and internal verification and review of environmental data.[6] To the extent implementation issues are discovered through assurance and subsequently resolved by

[6] For example, Apex Inc. provided assurance over Apple's Scopes 1 and 2 emissions. To conduct the assurance engagement, Apex undertook virtual remote site visits to Apple facilities in Israel and the U.S., conducted meetings with personnel from corporate offices in Cupertino, interviewed relevant personnel, reviewed internal and external documentary evidence, reviewed a sample of data against source data, and reviewed Apple's information system for collection, aggregation, analysis, and internal verification and review of environmental data. As another example, PwC provided assurance over Estée Lauder's Scopes 1 and 2 emissions, among others. PwC performed inquiries, conducted tests of mathematical accuracy of computations for sample data, read relevant policies, reviewed supporting documentation relating to the completeness and accuracy of data, and performed analytical procedures.

reporting firms, we should expect an improvement in carbon accounting quality post-assurance.

As discussed in the previous section, implementation issues could entail both errors and intentional misreporting. Our hypothesis is primarily based on assurors identifying implementation issues that do not necessarily involve managerial misreporting incentives. This is because the literature has not presented concrete evidence of managers manipulating emissions disclosure. Even in the earnings quality literature, the evidence of the relationship between managerial incentives and financial reporting quality is mixed (see, e.g., Armstrong et al., 2013). In addition, misreporting stems from incentives, which typically come from executive compensation contracts. As shown in Cohen et al. (2023a), despite an increasing trend, only 12% of U.S. public firms tie executive compensation to ESG metrics. Furthermore, common ESG metrics used in executive compensation include employee safety, security, and job satisfaction; only 1% of the firms link compensation to carbon emissions. Therefore, we leave it to future research to explore whether managerial incentives may drive manipulated reporting of emissions and whether the effect of assurance on carbon emissions is partially driven by these managerial incentives.[7]

3. Measuring Carbon Accounting Quality

3.1 Abnormal Emissions as a Measure of Carbon Accounting Quality

Understanding the role of assurance requires measures of carbon accounting quality, which we develop in this section. As discussed above, our first measure is based on the deviation of reported emissions from a model-based expected level. Carbon emissions modeling is a nascent field with works spanning environmental science, economics, finance, accounting, and business data vendors.[8] A common theme of those models is the fundamental idea that carbon emissions

[7] In other words, we posit that such a channel may exist, but we are agnostic about whether the incentive mechanism creates the implementation issues that we predict are lessened by assurance. To examine this mechanism in particular, future studies need to first establish a link between managerial incentives and carbon misreporting; then, studies can examine whether and when assurance can mitigate this effect.

[8] E.g., MSCI (2016), Goldhammer et al. (2017), Griffin et al. (2017), Nguyen et al. (2021), Downar et al. (2021), Cohen et al. (2023a, b), CDP (2022), Assael et al. (2023).

are a byproduct of a firm's production process; therefore, emissions are modeled as a function of predictors that capture a firm's production.

At the same time, those models differ in two interrelated dimensions. First, different studies use different sets of predictors, ranging from production-related firm characteristics (e.g., PP&E, firm sales) to financial statement ratios (ROA, book-to-market). Second, studies in economics, finance, accounting academia, and practice use simple calculations or linear models with a set of carefully selected determinants, whereas studies in engineering use machine learning, take a kitchen sink approach, and/or let the model select determinants. In addition, some studies tend to focus on narrow classes of firms, e.g., only electric utility firms, and use linear models with specific types of production, e.g., electricity produced at each utility's facility (e.g., Muller, 2021).

We tabulate the features of those models in Internet Appendix (IA) Table IA-2 where models can accommodate broad cross-sections of firms. To evaluate the relevance of determinants nominated by prior studies, we regress emission levels on each determinant and report the regression R^2. Consistent with emissions being a byproduct of production, variables related to fixed assets (e.g., asset tangibility, asset age, capital expenditure) and production output (e.g., inventory and cost of goods sold) have the highest explanatory power.

Given our survey of the literature, preliminary assessment of the determinants, and our discussions with environmental scientists, our modeling approach combines insights from the literature in both social science and engineering.[9] Specifically, we model emissions as a linear function of predictors as follows:

$$\text{Ln}(\mathit{Emissions}_{i,t,s}) = \beta_{industry(i),s}\ \mathit{Determinants}_{i,t,s} + \mathit{Year\ FE}_{industry(i),t,s} + \varepsilon_{i,t,s}, \tag{1}$$

[9] Similar to many social science studies, we model emissions as a linear function of predictors. We do not take a machine learning approach that includes all possible variables because we prefer to carefully select predictors rooted in the economics of production. However, we use the engineering literature to understand the emissions production process when selecting determinants.

where a firm i's Scope s emission level in year t is modeled as a function of determinants (more below) and year effects estimated for Scopes 1 and 2 separately. The economics of production likely differs by industry, so we allow the coefficients for each determinant to vary by industry.[10]

Our first set of determinants aims to capture production output. These determinants include firm sales, changes in inventory, and cost of goods sold. Because a firm's beginning inventory plus production minus products sold should equal its ending inventory, these determinants would allow us to back out production output. Because changes in inventory can be either positive or negative, we use the natural logarithms of contemporaneous and lagged inventory as opposed to the natural logarithm of changes in inventory, which is undefined when inventory change is negative.

Our second set of determinants intends to capture production technology. We include firm size and net property, plant, and equipment (PP&E) to proxy for the existence and scale of operating facilities, such as factories and machines that generate direct and indirect emissions (e.g., Nguyen et al., 2021). We also include fixed asset tangibility (net PP&E scaled by total assets) because firms with many fixed assets are likely to operate very differently compared to firms with many intangibles (Downar et al., 2021; Nguyen et al., 2021). In addition, we include capital expenditure and asset age (defined as accumulated depreciation, i.e., gross PP&E minus net PP&E, divided by annual depreciation) because production technology evolves, with more recent fixed assets more likely to be energy efficient and emitting less (e.g., Whitaker et al., 2012). We further consider the number of countries and states with operations (Goldhammer et al., 2017). *Ceteris paribus*, we expect centrality of production to generate economies of scale in production and

[10] By allowing for industry-specific coefficients for each determinant, our model is analogous to estimating panel regressions for all firms in each industry. Perhaps the most distinctive difference between our model and the abnormal accruals models is that our model is estimated for all firms within each industry, whereas abnormal accruals are estimated for each industry-year pair. Because carbon emissions disclosures are voluntary, estimating Eq. (1) for each industry-year pair will result in many industry-years with insufficient observations if requiring a minimum number of observations, as done in papers estimating the abnormal accruals, especially in early years. We acknowledge this is a limitation. With the potential for mandatory disclosure in the future, subsequent research could introduce separate coefficients for each industry-year pair.

reduce emissions; centrality also decreases the need for intra-firm transportation, reducing emissions. Last, we include the number of employees. For new economy firms, Scope 2 emissions come primarily from office space utilities, which are determined primarily by employee base.

Aside from production output and technology, we include historical emissions in our model. This is a common element in proprietary models built by environmental experts for sustainability data vendors (e.g., MSCI, Refinitiv); those models use historical emissions to extrapolate current emissions. Conceptually, when including lagged emissions, the residual term is then the amount of *innovation* in emission levels that cannot be explained by a firm's production activities. Therefore, the residual is more likely to capture (innovations in) errors and issues in the carbon accounting system. Because we do not estimate Eq. (1) for each industry-year pair, controlling for lagged emissions also accounts for firms that may have different production technologies and, hence, emissions.[11] To further account for technology, we include industry-year fixed effects.

We use predicted residuals from this model of firm-year-scope-level emissions, i.e., the predicted value of ε or modeled versus reported emissions, as our measure of carbon accounting quality. Because we test predictions regarding implementation issues in carbon accounting systems that could symmetrically affect emissions measurement, we use the absolute value as our measure of carbon accounting quality in our subsequent tests. We call this variable *Abnormal Emissions*.

Model misspecification may affect our inferences. We are conceptually interested in the relation between errors in emissions reporting coming from a very well-specified model of emissions and assurance, but via our implementation, we regress the predicted value of ε from a model of production on assurance. To the extent that there is model misspecification and the misspecification relates to assurance, our estimated coefficient of interest could be biased. To correctly measure an effect of assurance on carbon accounting quality, we need to at least reduce

[11] In a robustness test, we estimate an alternative version of the model where we drop lagged emissions and instead use changes in total emissions as the dependent variable. See Table IA-4.

the possibility that we have a systematically misspecified model of emissions or utilize potentially exogenous sources of variation in assurance. For the former, we evaluate our emissions model on several dimensions, such as explanatory power, in Section 4.3. We also use alternative, non-regression-based measures of carbon accounting quality. For the latter, we exploit a shock of mandatory assurance in the E.U., where we believe the endogeneity problem to be less severe.

3.2 Alternative Measures of Carbon Accounting Quality

A key advantage of our regression-based measure is that it can be estimated for any firm with disclosure of emissions, which could be particularly valuable for firms without CDP disclosure and especially when carbon emissions disclosure is mandated. However, given the model misspecification concern discussed above, we triangulate our inference using two alternative measures of carbon accounting quality from the CDP questionnaires.

In some years of our sample period, CDP asks reporting firms to identify how uncertain they are about the reported emission figures, with options including below 2%, 2-5%, 5-10%, 10-20%, all the way up to 100%. This measure captures a firm's *perception* of its carbon accounting quality. To the extent a manager is on average truth-telling, smaller managerial uncertainty should imply high carbon accounting quality.

In addition, a well-functioning carbon accounting system should be able to generate carbon disclosures more efficiently and timely. In the accounting quality literature, studies use the time difference between the fiscal year end date and the earnings announcement (or annual report filing) date as a measure of financial reporting system quality (e.g., Heitzman and Huang, 2019). Similarly, we use the difference between the CDP questionnaire release date and firm response submission date (*Filing Lag*) as an alternative measure of carbon accounting system efficiency. We note that Scopes 1 and 2 emissions are just a very small part of the many questions included in CDP questionnaires. Therefore, this measure is indirect and not scope-specific.

4. Data, Descriptive Statistics, and Measure Validation

4.1 Data

To construct the sample, we combine carbon emissions data from Trucost, assurance data from Gipper et al. (2023) and CDP, and firm fundamentals from Compustat. Trucost carbon emissions data include both disclosed emissions from reporting firms as well as estimated emissions from non-disclosing firms. Because we are interested in the quality of firms' carbon emissions disclosures, we only consider firm-years with disclosed emissions. Data on assurance used in this study come from Gipper et al. (2023) and CDP. Gipper et al. (2023) compiled a comprehensive set of ESG reports issued by U.S. public firms and manually collected detailed information on disclosure and various dimensions of assurance (e.g., assuror identity, level of assurance) at the metric level (hereafter, "GRS data"). This is the only data source that tracks the various characteristics of ESG assurance by metric, whereas major commercial databases (e.g., Corporate Register, Refinitiv) treat ESG assurance as binary, measure assurance at the firm-year level regardless of which metric is assured, and omit as many as 50% of the assurance instances. We supplement GRS data with CDP. Sometimes, firms report their GHG disclosure and assurance to the CDP but do not issue standalone ESG reports. By combining GRS data with CDP data, we reduce the measurement error in reporting and assurance.[12]

In our analyses, we use several samples. For our main analyses focusing on U.S. firms, our sample is a panel spanning 2010 to 2020, covering 987 firms. The number of observations varies for Scopes 1 and 2 emissions because not all firms disclosing Scope 1 emissions also disclose emissions from Scopes 2 or 3. We use this sample for our main, robustness, and cross-sectional analyses. This sample reflects the intersection of Compustat, CRSP, and Trucost with assurance coverage from GRS and/or CDP data. It includes all U.S. public firms with non-missing controls

[12] To the extent we do not capture all instances of assurance, our setting biases against us finding a result because assurance that we failed to measure would be pooled with non-assurance firms.

and data to estimate *Abnormal Emissions*. We also perform several tests using additional data from the CDP questionnaires.

4.2 Descriptive Statistics

Table 1 presents descriptive statistics. Notably, the conditional rate of assurance in our sample is higher than Gipper et al. (2023). This is because the numbers in Gipper et al. (2023) include assurances in ESG reports, whereas firms can disclose (and assure) emissions in other venues, such as through CDP and in annual reports. However, the numbers are in line with industry reports. For example, the Center for Audit Quality incorporates CDP data and reports an assurance rate of 56.6% for S&P 500 firms in 2020. Given that our sample includes earlier years, where assurance was less common, we believe our reported statistics are reasonable.

Table 1 also reports various indicators of the carbon accounting system, which we derive from the CDP questionnaire data. The sample size varies by variable because the CDP questionnaire varies by year and some questions are only included in select years. In addition, these questions are conditional on firms disclosing an emissions number. Because firms are more likely to disclose Scope 1 emissions relative to Scope 2 emissions, we see larger sample sizes for Scope 1 emissions. On average, assurors identify two categories of issues. The probability of omitting sources of emissions is 0.45 and 0.37 for Scopes 1 and 2, respectively, suggesting that reporting issues are pervasive.

4.3 Validating Abnormal Emissions as a Measure of Carbon Accounting Quality

In addition to using alternative measures, we strengthen our inference by validating our *Abnormal Emissions* measure in four ways. First, a criticism of abnormal accruals models is that determinants only explain roughly 30% of the total variation in working capital changes (e.g., McNichols, 2002). For carbon emissions, our determinants explain 96.2% and 91.0% of the total variation in emission levels for both Scopes 1 and 2.[13] Given the limited residual variation, we believe the measurement error problem is not nearly as severe as for abnormal accruals models.

Second, as firms gain more experience in reporting carbon emissions, we should expect improvements in carbon accounting quality within-firm over time, especially due to the increasing focus on emissions during the decade. We regress *Abnormal Emissions* on indicators for each year, with 2010 as the base year and, hence, omitted. In generating these figures, we include firm fixed effects to focus the analysis on the time series and mitigate the impact coming from the change in the mix of disclosing firms over the decade. Figure 1 depicts the time-series variation in abnormal emissions. We find decreasing trends for both Scope 1 and Scope 2 abnormal emissions, suggesting that firms improve their carbon accounting quality over time.

Third, we relate the abnormal emissions measure to the managerial uncertainty measure of carbon accounting quality based on CDP questionnaires. As we show in Table 2 Panel A, we find a strong relation between *Abnormal Emissions* and *Manager Uncertainty* for both Scopes 1 and 2 emissions. To better understand the association, we examine binned scatterplots. Specifically, we sort *Abnormal Emissions* into ten equal-sized bins. For each bin, we calculate the mean of *Abnormal Emissions* and *Manager Uncertainty*, and then plot the means for each of the ten bins. As we show in Figure 2, the positive associations documented in Table 2 are mostly linear, i.e., the higher the *Abnormal Emissions*, the higher the *Manager Uncertainty*.

Lastly, we relate the abnormal emissions measure to disclosure scores from Bloomberg and Refinitiv. Refinitiv has a score that captures a firm's overall ESG reporting practice, and Bloomberg has a score that specifically captures a firm's environmental disclosures. For both rating agencies, a higher score indicates better disclosure practices. Because firms with high-quality carbon accounting likely have more robust ESG reporting, we expect *Abnormal Emissions*

---

[13] In our carbon emissions model, we allow the coefficient of each determinant to vary by industry. For parsimony, in Table IA-3, we present a version of our carbon emissions model without industry-level coefficients.

to be negatively correlated with ESG disclosure scores. Consistent with this, we find in Table 2 Panel B that firms with higher Bloomberg environmental disclosure scores and higher overall Refinitiv disclosure scores have higher carbon accounting quality for both Scopes 1 and 2 (i.e., lower *Abnormal Emissions*). Together, the findings in Figures 1-2 and Table 2 provide us with confidence that our *Abnormal Emissions* measure reasonably captures carbon accounting quality.

5. Research Design

To investigate the association between assurance and carbon accounting quality, we estimate the following regression model:

$$CAQ_{i,t,s} = \beta \ Assurance_{i,t,s} + \text{Controls} + \text{Fixed Effects} + \varepsilon_{i,t,s} \qquad (2)$$

In Eq. (2), *CAQ* is a measure of carbon accounting quality for firm *i*'s Scope *s* emissions in year *t*. Our main measure of *CAQ* is *Abnormal Emissions*. The larger the variable, the lower the carbon accounting quality. We triangulate our inference using manager's assessment of emissions reporting uncertainty (*Manager Uncertainty*) and CDP questionnaire filing lags (*Filing Lag*) as two alternative measures of *CAQ*. Our independent variable of interest, *Assurance*, is an indicator variable equal to one if a firm's carbon emissions for Scope *s* are assured in year *t*. For all three measures, smaller values indicate higher carbon accounting quality. Therefore, we predict β<0. Although there is no theoretical guidance for the selection of controls, we consider common firm characteristics, including firm size, ROA, book-to-market, leverage, and an indicator for loss firms (Gipper et al., 2023). The purpose of including these firm characteristics is to mitigate the concern that firm characteristics jointly determine assurance and carbon accounting quality. For example, larger firms may have the resources to set up a robust carbon accounting system and engage an assuror. Controlling for firm size allows us to rule out this possibility. Eq. (2) includes industry and year fixed effects. Year fixed effects remove time trends in carbon accounting quality. Inherently different production functions across industries could lead to heterogeneous measurement problems in emissions, resulting in differences in *CAQ*. Industry fixed effects remove the influence of these differences, along with other industry-invariant confounding factors. All continuous variables are winsorized at the 1% and 99% levels. Throughout the study, we base our inferences on standard errors clustered by firm.

6. Results

6.1 Main Result

We report our main findings in Table 3 Panel A. In columns (1) and (3), we report the univariate correlation between assurance and carbon accounting quality for Scopes 1 and 2, respectively. Note that both *Assurance* and *Abnormal Emissions* are measured at the firm-year-scope level. Because *Abnormal Emissions* inversely captures carbon accounting quality, the negative coefficients are consistent with third-party assurance *improving* carbon accounting quality. In columns (2) and (4), we include industry and year fixed effects, as well as common firm characteristics as control variables. Those controls and fixed effects partially mitigate the concern that firm and/or industry characteristics jointly determine improvement in carbon accounting quality and the decision to assure emissions metrics. For example, the coefficient on firm size is negative, suggesting that larger firms have higher carbon accounting quality. This finding is in line with larger firms having better financial reporting quality (Dechow et al., 2010). The coefficient on *Assurance* continues to be negatively significant. More importantly, the economic magnitudes are sizable. Assurance increases the quality of Scopes 1 and 2 emissions disclosure by 26 and 34% relative to the sample mean, respectively.

To mitigate the concern that our finding is driven by the misspecification of the emissions model, we show in Table IA-4 that our findings are robust to modifications of the research design, including the emissions model. First, we allow each determinant of the emissions model to have separate coefficients for each year or each industry and half-decade combination (whereas in the baseline model, each determinant has a separate coefficient for each industry), and these alternatives do not change our inference. Second, our finding is robust to using changes in emissions as the dependent variable, an approached used in some studies (e.g., Cohen et al., 2023b). Lastly, we perform several tests to address concerns raised by Chen et al. (2018, 2023) that inferences can be incorrect when the residual from a regression model is used as the dependent variable. For example, our result is robust to including all determinants from the carbon emissions model as additional controls, adding assurance as an additional determinant in the carbon emissions model, and estimating quantile regressions with the level of emissions.[14]

In Panel B, we examine the robustness of our finding using two alternative measures of carbon accounting quality. First, we focus on managers' confidence in their carbon accounting systems. In some years, firms are asked to assess how certain they are about their firm's scope-specific emissions disclosure. In columns (1) and (2) of Panel B, we find assurance reduces firms' uncertainty about emissions. Next, we examine the efficiency of the carbon accounting system. A more efficient carbon accounting system should be able to generate carbon disclosures more efficiently and timely. We use the difference between the CDP questionnaire release date and firm response submission date (*Filing Lag*) as a measure of carbon accounting system efficiency. In columns (3) and (4) of Panel B, we show that assurance is associated with shorter filing lags. In Table IA-4, our findings are similar if we use the first factor of a principal component analysis of *Abnormal Emissions*, *Manager Uncertainty*, and *Filing Lag* as the dependent variable. Together, the findings in Table 3 present strong evidence that assurance relates to carbon accounting quality and is consistent with assurance *increasing* carbon accounting quality.

6.2 Addressing Alternative Explanations

Absent exogenous variation, our main finding could be subject to alternative explanations. One alternative is that the documented association is driven by a latent variable, such as firm type.

---

[14] Because our analyses use a two-stage method, we bootstrap with firm-level blocks over both the emissions model and the main test results to correct for standard error measurement and find that, if anything, the standard errors are slightly smaller with such a correction. For instance, firm-level clustering generates a standard error estimate of 0.0156 for the Scope 1 variable of interest while the two-step bootstrapped standard errors with firm clustering generate an estimate—with 1000 replications—of 0.0150. We use the more conservative standard error estimates.

For example, a transparent firm simultaneously chooses assurance and a high-quality carbon accounting system that produces high-quality carbon emissions numbers. We address this concern in several ways. First, we use firm fixed effects to remove variation driven by firm type. As we show in Panel A of Table 4 columns (1) and (4), the result is robust to within-firm variation.

Second, we explore changes in abnormal emissions following first-time assurance adoptions. To do so, we create an indicator equal to one if an observation is in or after the first year with assurance for each scope. We regress *Abnormal Emissions* on this indicator, with either industry or firm fixed effects, akin to a staggered difference-in-differences design. As we show in columns (2), (3), (5), and (6), the results continue to indicate a negative association between *Abnormal Emissions* and *Assurance*. Comparing the coefficients in, e.g., columns (1) and (3), we find that *Post Assurance* appears to have a larger effect than *Assurance*, indicating that annual assurance might not be necessary to achieve the measurement benefits from verification.

Third, we show that the association between carbon assurance and carbon accounting quality is scope-specific. To do so, we include assurance indicators for both Scope 1 and Scope 2 on the right-hand side of Eq. (2) and study their associations with abnormal emissions. Because there is variation in assurance across scopes, including both indicators in the equation is essentially a horserace. As we show in column (1), when *Abnormal Emissions* is measured for Scope 1 emissions, the coefficient on *Scope 1 Assurance* (i.e., assurance over Scope 1 emissions) has a larger economic magnitude and is more statistically significant relative to the coefficient on *Scope 2 Assurance* (i.e., assurance over Scope 2 emissions). Similarly, when *Abnormal Emissions* is measured for Scope 2 emissions in column (2), the coefficient on *Scope 2 Assurance* has a larger magnitude, a negative association, and significance, unlike the coefficient on *Scope 1 Assurance*.

Fourth, we show that assurance over emissions metrics does not statistically correlate with measures of financial reporting quality, including an indicator of whether a firm just meets or beats its earnings target (Burgstahler and Dichev, 1997) and abnormal accruals per McNichols (2002).

If our main result is driven by overall firm transparency or financial reporting quality, we would expect a positive association between carbon assurance and financial reporting quality. However, as we show in columns (3)-(6), we do not find carbon emissions assurance to be correlated with measures of financial reporting quality, suggesting that our finding is unique to carbon reporting.

Lastly, we explore how assurance shapes carbon accounting quality in event time. We identify the first year a firm decides to assure its Scopes 1 or 2 emissions and regress *Abnormal Emissions* on indicators for each year relative to the first year of assurance adoption. We include controls in Eq. (2) and firm fixed effects to mitigate the influence of confounders. As shown in Figure 3, in the pre-period, the coefficients are small, statistically insignificant, and exhibit no apparent trends for Scope 1 emissions. For Scope 2 emissions, all coefficients except for one are statistically insignificant in the pre-period, with a slight downward trend from t = -2. One possibility is that firms assure Scope 1 emissions before Scope 2 emissions. Because we identify the first-time Scope 2 assurance year as the first year of treatment and firms likely use the same carbon accounting system for Scopes 1 and 2 emissions, the Scope 2 pre-trend likely reflects the spillover effect from Scope 1 assurance already in place. In the post-period, for both scopes, all coefficients are negative and statistically significant, suggesting a decrease in *Abnormal Emissions*. Importantly, there is a significant "jump" in the coefficient magnitude in the first year of treatment (from t = -1 to t = 0). This trend is consistent with assurance improving carbon accounting quality, as opposed to an alternative "assurance readiness" explanation in which firms obtain assurance when their carbon accounting quality is already high (or has recently improved) and assurance serves as a signaling device. Further, the coefficient magnitude becomes more negative over time, suggesting that firms continue to work with assurors to improve their carbon accounting quality.

6.3 Cross-sectional Variation in Assurance Quality

If assurance indeed improves carbon accounting quality, our main finding should be stronger when assurance is thorough and high-quality. We leverage our detailed assurance data and develop two measures that capture the effort that goes into assurance. Our first indicator of assurance quality focuses on the thoroughness of assurance within a scope—the contractual level of assurance. When firms hire an assurance provider, the engagement contract would specify whether the outcome of the assurance engagement is reasonable assurance (like for audits of annual financial statements) or limited assurance, which is less involved in terms of the work performed by the assuror (like for reviews of quarterly financial statements). Because reasonable assurance is much more thorough than limited assurance, we expect the association between assurance and abnormal emissions to be stronger for reasonable assurance.

Our second measure of assurance quality is whether a firm assures all three scopes of reported emissions. To assure all three emissions scopes, an assurance provider inherently reviews more data and conducts more thorough investigations of the internal carbon accounting system. In addition, because firms choose reporting boundaries that separate Scopes 1, 2, and 3 emissions, engaging an assuror for all three scopes of emissions allows the assuror to more accurately identify reporting boundaries and therefore reduce omissions and misclassifications.

To perform the cross-sectional analysis, we modify Eq. (2) by partitioning the scope-specific *Assurance* indicator into non-overlapping variables as follows.

$$CAQ_{i,t,s} = \sum \beta_N \, Assurance\ Characteristics_{i,t,s} + \text{Controls} + \text{Fixed Effects} + \varepsilon_{i,t,s}. \qquad (3)$$

Assurance Characteristics are partitioning variables, which vary by analysis. For example, when we investigate whether the association between assurance and carbon accounting quality varies by the contractual level of assurance, *Reasonable Assurance* would equal *Assurance* when the contractual level of assurance is reasonable, and zero otherwise, and *Limited Assurance* would equal *Assurance* when the contractual level of assurance is not reasonable (i.e., limited assurance), and zero otherwise. The coefficients on *Limited Assurance* and *Reasonable Assurance* capture the effect of reasonable assurance and limited assurance on carbon accounting quality, respectively. Together, *Reasonable Assurance* and *Limited Assurance* span *Assurance*. This approach has been frequently used in prior studies (e.g., Christensen et al., 2016; Zhu, 2019; Jayaraman and Wu, 2019). An advantage of this design is that it allows us to directly estimate the effect of assurance for the two partitioning groups. To statistically compare the heterogeneous effects, we perform a Wald test of coefficient equality for our partitioning variables.

We report the findings in Table 5, with odd and even columns reporting the results for Scopes 1 and 2, respectively. Columns (1) and (2) focus on whether the effect of assurance is heterogeneous in the contractual level of assurance. For both scopes, the coefficient on *Reasonable Assurance* is larger, meaning that relative to limited assurance, reasonable assurance further reduces the level of abnormal emissions. Although the estimated magnitude for reasonable assurance is larger, the difference is not significant at conventional levels.[15] In columns (3) and (4), we examine whether the effect of assurance varies by the pervasiveness of assurance. Consistent with more thorough assurance resulting in higher reporting quality, firms that assure all three scopes of assurance have lower levels of abnormal emissions. Notably, the difference in coefficients is statistically significant at conventional levels for Scope 2 emissions.

Next, we combine the two dimensions of assurance quality and partition *Assurance* into four non-overlapping cases: (1) limited assurance with not all three scopes assured, (2) limited assurance with all three scopes assured, (3) reasonable assurance with not all three scopes assured, and (4) reasonable assurance with all three scopes assured. We expect scenario (4) to have the highest assurance quality and (1) to have the lowest assurance quality. We do not have strong expectations with regard to the relative assurance quality of (2) and (3). Consistent with the prediction, in columns (5) and (6), we find the coefficient magnitude of *Assurance* to be the largest for (4) and smallest for (1). In terms of the magnitude difference, having reasonable assurance and

[15] Anecdotally, when a firm starts to seek assurance, the reporting firm and the assuror typically start with limited assurance and sometimes set the goal to eventually converge to reasonable assurance. This upward trajectory implies that within the category of limited assurance, there is likely significant heterogeneity in reporting quality, with some closer to reasonable assurance and some barely meeting the limited assurance standard. The pooling of reporting types within limited assurance could explain why the magnitude difference is not significant at conventional levels.

all three scopes assured results in a 32% and 49% increase in carbon accounting quality relative to having limited assurance and only some scopes assured. In aggregate, these analyses provide evidence broadly consistent with thorough assurance improving carbon accounting quality.[16]

6.4 Mechanism

Our next set of tests focuses on *how* assurance improves carbon accounting quality. As discussed in Section 2.2, this process likely involves assurors identifying and fixing various implementation issues. If so, we expect assurance to be associated with various *indicators* of the carbon accounting system. To test this hypothesis, we turn to CDP data and develop measures based on firms' responses to CDP questionnaires.[17]

We first focus on indicators that are inputs to generating carbon accounting numbers. In the assurance process, assurors will likely identify various issues related to carbon accounting, including using incorrect conversion factors and failing to account for energy use in a facility, among others. Therefore, we expect *Assurance* to be positively associated with the probability of issues being identified (*Issues Identified*), and consequently, negatively associated with the number of emissions sources being omitted when reporting emissions (*Omissions*). As we show in columns (1) through (4) of Table 6, assurance increases issues identified and decreases omissions. Interestingly, the magnitude of the association between *Assurance* and *Omissions* is larger for Scope 2 emissions. One possibility is that emission sources for Scope 1 are easier to comprehensively identify in the first place, limiting the incremental role of assurance.

Once implementation issues in the carbon accounting system are identified, assurors would

[16] In Table IA-5, we perform two additional cross-sectional analyses. First, we examine whether the effect is heterogeneous in whether the assurance is provided by a financial auditor vs. non-financial assuror. We find no meaningful difference between the two. Second, we conduct a descriptive split which shows carbon accounting quality changes when obtaining assurance depending on whether firms have emissions reduction targets. We find a stronger effect for firms with existing targets.

[17] For Table 6, sample size varies by dependent variable. This is because CDP questionnaires vary by year and some questions are not consistently asked. In addition, although CDP tracks when firms submit their responses, this information is not available to us for some years in our sample period.

communicate the issues to the reporting firm. To the extent firms take remedial actions, we would expect historically reported numbers to be revised. For example, if an assuror identified that a reporting firm's emissions number fails to account for emissions from an overseas facility, the firm could re-calculate and update the number. More formally, the GHG Protocol explicitly states that revisions are required when (1) there are significant changes in calculation methodology, which typically results from improved emission factors and/or improved activity data and (2) significant errors or smaller errors that are collectively significant are discovered.[18] Consistent with our prediction and prior literature, the results in columns (5) and (6) show that assurance is associated with an increased probability of firms revising their historical emissions figures.

Taken together, the results in Tables 3 and 6 describe how assurance improves carbon accounting quality. First, assurors identify issues and communicate them to the reporting firm, and there is a decline in reporting omissions, consistent with an improved and more complete carbon accounting system. Then, firms take remedial actions and improve their carbon reporting. As a result, managers believe that their reported emissions figures are more precise, and firms can assemble and file emissions information with CDP in a timelier manner.[19]

6.5 International Evidence

Thus far, our analysis has focused on public firms in the U.S. Are the findings generalizable to other countries? Using disclosed emissions from Trucost and firm fundamentals from Compustat Global, we estimate abnormal emissions for 42 countries with available data. We report

---

[18] This argument is similar to financial auditors identifying internal control weaknesses after SOX, resulting in more frequent restatements and disclosure of control weaknesses (e.g., Ge and McVay, 2005; Doyle et al., 2007; Rice and Weber, 2012). However, in emissions reporting, measurement standards and assurance practices are not as consistent and well understood as in financial reporting, leading to higher revision rates than financial statement restatements (e.g., Michelon et al., 2019). We use different terminology, "revisions," reflecting that the practices are different, such as revising emission numbers from prior years for methodological changes—a practice that is specifically disallowed in financial reporting. Moreover, assurors may signal quality through emissions revisions (Michelon et al., 2019), but undiscovered financial misstatements later revealed through a restatement are a signal of low audit quality (DeFond and Zhang, 2014).

[19] The fact that assurance increases the number of issues identified helps us rule out the alternative story that firms assure their emission numbers when their carbon accounting quality is high or improving. If the alternative story is true, we should observe a decrease in the number of issues identified with the carbon accounting system.

sample composition by country in Panel A of Table 7. Outside of the U.S., Japanese and U.K. firms comprise more than 30% of the total sample. This is unsurprising, as both countries have many public firms and reporting mandates for greenhouse gas emissions. Panel B of Table 7 provides descriptive statistics for this international sample. We identify assurance using solely CDP questionnaires because the GRS sample does not include international firms. Gipper et al. (2023) note that because emissions (and other ESG) assurance is not consistently structured within firms' reports, or even disclosed, data sources likely undermeasure assurance. Thus, it is possible that some firms in our sample have obtained assurance over emissions metrics but do not have CDP coverage; this would bias *against* us finding a result.

For the international sample, we modify our baseline abnormal emissions model by including country fixed effects, recognizing the heterogeneity in production technologies across countries. For example, a firm from a less developed country might use old production technologies that generate more emissions than a firm from an advanced economy with state-of-the-art equipment. Because our measure of production centrality (the number of states and countries with production facilities) is unavailable for non-U.S. countries, we drop this determinant for our international analysis. As shown in Table 7 Panel C, we continue to find a positive relation between assurance and carbon accounting quality. In terms of the economic magnitude, columns (1) and (4) suggest that assurance reduces abnormal emissions by 15 and 11% relative to the sample mean for Scopes 1 and 2, respectively. In addition, columns (3) and (6) suggest that after a firm starts to assure its emissions, abnormal emissions reduce by 18 and 14% relative to the sample mean for Scopes 1 and 2, respectively. Together, the results in Table 7 show that our main finding is not specific to U.S. firms.

6.6 Shock to ESG Assurance: E.U.'s Non-Financial Reporting Directive

In the last analysis, we utilize the implementation of the E.U.'s Non-Financial Reporting Directive 2014/95/EU (NFRD) to further examine how assurance shapes carbon accounting quality. Adopted in 2014 and effective in all E.U. member states in 2018, NFRD requires public-interest entities (such as publicly listed companies, banks, or insurance companies) exceeding 500 employees on average to report on a variety of ESG issues, including environmental matters. If a firm considers climate as a material issue, the firm is required to disclose climate information. As a part of NFRD, all E.U. countries have requirements for the statutory auditor to check whether non-financial information has been provided. On top of that, 11 countries have additional requirements for the auditor to check the consistency of non-financial information with the financial statements. Finally, France, Italy, and Spain require mandatory independent assurance over the non-financial information from this directive.[20]

We leverage this institutional setting and examine whether mandatory assurance over non-financial information, which likely includes GHG emissions for many firms, increases carbon accounting quality. To do so, we exploit within-E.U. variation in assurance and estimate difference-in-differences regressions as follows:

$$CAQ_{i,t,s} = \beta_1 \times Mandatory \times Post + \text{Controls} + \text{Firm, Year FE} + \varepsilon_{i,t,s}. \qquad (4)$$

In Eq. (4), the dependent variable is *Abnormal Emissions* for firm *i*, Scope *s*, in year *t*. *Mandatory* is an indicator variable measured at the country level and equals 1 if a firm is in France, Italy, or Spain, and 0 otherwise. *Post* equals 1 if the observation is after NFRD implementation, and 0 otherwise. We use the same set of control variables. In addition, we include firm and year fixed effects, which subsume *Mandatory* and *Post*, respectively. The use of firm fixed effects is particularly important in this setting. Given that NFRD is not a regulation specific to greenhouse gas emissions, firm fixed effects allow us to abstract away from firm-invariant confounders, e.g., the types of non-financial information disclosed. In addition, we use a short event window from

[20] In untabulated tests, we confirm that firms in these countries have relatively higher assurance rates after the directive. The increase in assurance rates is not specific to any one country, alleviating the possibility that prior rules or early adoption concentrated in a country reduces the usefulness of this setting as a shock to emissions assurance.

2017-2020, two years before and after NFRD implementation, which sharpens our identification. For this analysis, we consider two sets of control firms. The first set consists of all firms in E.U. countries without an assurance mandate, and the second set consists of firms in E.U. countries that require auditor consistency checks. Our coefficient of interest is on *Mandatory* × *Post* (i.e., β_1), which we predict to be negative. As we show in columns (1)-(4) of Table 8, mandatory assurance improves the reporting quality for Scope 1 emissions. For Scope 2 emissions, firms in countries with the assurance mandate still experience lower abnormal emissions post-NFRD; however, the magnitudes are smaller. In untabulated tests, we find that lower abnormal emissions post-NFRD is not specific to any single country with the assurance mandate.

To provide additional credibility, we perform a falsification test. Recall that 11 countries require auditors to check the consistency of non-financial information with the financial statements but do not require mandatory assurance. If our documented effect is not assurance-specific, then we would expect a similar result for those 11 countries. In columns (5) and (6), we perform a similar DiD analysis. In this falsification test, the treatment group includes firms from the 11 countries and the control group includes firms from the remaining E.U. member states, excluding France, Italy, and Spain. Consistent with our conjecture, we find no evidence that "consistency checks" improve carbon accounting quality. Overall, the evidence in Table 8 supports our inference that assurance shapes carbon accounting quality.

7. Conclusion

This study examines the role of third-party assurance in shaping carbon accounting quality. We propose a conceptual framework of carbon accounting quality and introduce a scope-specific measure inspired by abnormal accruals models. We document a positive association between assurance and carbon accounting quality for both U.S. and non-U.S. countries. This relation is stronger when assurance is more thorough. We also document how assurance improves carbon accounting quality: first, assurors identify issues in the carbon accounting system and communicate them to the firm; subsequently, firms take remedial actions, resulting in updated disclosures. Using the implementation of NFRD in the E.U. as an alternative setting, we show that mandatory assurance improves carbon accounting quality.

Our study contributes to the ESG reporting quality literature by introducing a novel measure of emissions reporting quality. Notably, this measure can be derived for any firm disclosing emissions data and is scope-specific. We believe this measure could be useful for academic researchers and stakeholders attempting to assess carbon disclosure quality. Our study also contributes to the ESG assurance literature by documenting the economic effects of assurance. Gipper et al. (2023) document significant heterogeneity in ESG assurance and the pervasive use of limited assurance in ESG reporting, casting doubt on the usefulness of ESG assurance. On the contrary, our findings suggest that even limited assurance can shape carbon accounting quality; this is the level of assurance included in the SEC's proposed rules for climate-related disclosures.

We acknowledge that our regression-based measure of abnormal emissions is subject to model misspecification and hence measurement error. Starting from Jones (1991), there have been volumes of studies attempting to develop and evaluate measures of earnings quality for the subsequent 30 years, and the literature is still expanding (Dechow et al., 1995; Dechow and Dichev, 2002; McNichols, 2002; Kothari et al., 2005; Breuer and Schütt, 2023). We do not believe that our study can develop a perfect measure of carbon accounting quality, but we hope that the methodology proposed in the paper could be a useful starting point. Future research could either build on our models or devise alternative measures of carbon accounting quality. With respect to assurance, our study does not establish a causal relation between assurance and carbon accounting quality, but our tests show that carbon accounting quality is improving concurrently with assurance in ways that are consistent with causality (e.g., improvements at the scope level and following non-financial information-specific assurance mandates). Given the looming assurance mandates in the U.S. and E.U., we believe this could be another fruitful area for future research.

References

Armstrong, Christopher S., David F. Larcker, Gaizka Ormazabal, and Daniel J. Taylor. "The Relation between Equity Incentives and Misreporting: The Role of Risk-Taking Incentives." *Journal of Financial Economics* 109, no.2 (2013): 327-350.

Assael, Jérémi, Thibaut Heurtebize, Laurent Carlier, and François Soupé. "Greenhouse Gases Emissions: Estimating Corporate Non-Reported Emissions Using Interpretable Machine Learning." *Sustainability* 15, no. 4 (2023): 3391.

Ballou, Brian, Po-Chang Chen, Jonathan H. Grenier, and Dan L. Heitger. "Corporate social responsibility assurance and reporting quality: Evidence from restatements." *Journal of Accounting and Public Policy* 37, no. 2 (2018): 167-188.

Barth, Mary E., Steven F. Cahan, Li Chen, and Elmar R. Venter. "The economic consequences associated with integrated report quality: Capital market and real effects." *Accounting, Organizations, and Society* 62 (2017): 43-64.

Burgstahler, David, and Ilia Dichev. "Earnings management to avoid earnings decreases and losses." *Journal of Accounting and Economics* 24, no. 1 (1997): 99-126.

Bolton, Patrick and Marcin Kacperczyk. "Do investors care about carbon risk?" *Journal of Finance Economics* 142, no.2 (2021): 517-549.

Bourveau, Thomas, Janja Brendel, and Jordan Schoenfeld. "Decentralized finance (DeFi) Assurance: Audit Adoption and Capital Market Effects." Working Paper. Available at SSRN 4457936. (2023a).

Bourveau, Thomas, Matthias Breuer, Jeroen Koenraadt, and Robert Stoumbos. "Public Company Auditing Around the Securities Exchange Act." Working Paper. Available at SSRN 3837593. (2023b).

Breuer, Matthias and Harm H. Schütt. "Accounting for uncertainty: an application of Bayesian methods to accruals models." *Review of Accounting Studies* 28, no.2 (2023): 726-768.

CDP. "CDP Full GHG Emissions Dataset: Technical Annex III: Statistical Framework." (2022).

Center for Audit Quality. 2023. "S&P 500 ESG Reporting and Assurance Analysis." Available at https://www.thecaq.org/sp-500-and-esg-reporting

Chen, Wei, Paul Hribar, and Sam Melessa. "Incorrect Inferences When Using Residuals as Dependent Variables." *Journal of Accounting Research* 56, no.3 (2018): 751-796.

Chen, Wei, Paul Hribar, and Sam Melessa. "On the Use of Absolute Residuals as Dependent Variables in Accounting Research." Working Paper. Available at SSRN 4572590. (2023).

Christensen, Hans B., Luzi Hail, and Christian Leuz. "Capital-Market Effects of Securities Regulation: Prior Conditions, Implementation, and Enforcement." *The Review of Financial Studies* 29, no.11 (2016): 2885-2924.

Cohen, Shira, Igor Kadach, Gaizka Ormazabal, and Stefan Reichelstein. "Executive Compensation Tied to ESG Performance: International Evidence." *Journal of Accounting Research* 61, no.3 (2023a): 805-853.

Cohen, Shira, Igor Kadach, and Gaizka Ormazabal. "Institutional investors, climate disclosure, and carbon emissions." *Journal of Accounting and Economics* 76, no.2 (2023b).

Dechow, Patricia M., Richard G. Sloan, and Amy P. Sweeney. "Detecting Earnings Management." *The Accounting Review* 70, no.2 (1995): 193-225.

Dechow, Patricia M. and Ilia D. Dichev. "The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors." *The Accounting Review* 77, (2002): 35-59.

Dechow, Patricia M., Weili Ge, and Catherine M. Schrand. “Understanding Earnings Quality: A Review of the Proxies, Their Determinants and Their Consequences.” *Journal of Accounting and Economics* 50, no. 2-3 (2010): 344-401.

DeFond, Mark, and Jieying Zhang. “A review of archival auditing research.” *Journal of Accounting and Economics* 58, no. 2-3 (2014): 275-326.

Doyle, Jeffrey, Weili Ge, and Sarah McVay. “Determinants of weaknesses in internal control over financial reporting.” *Journal of Accounting and Economics* 44, no.1-2 (2007): 193-223.

Downar, Benedikt, Jürgen Erstberger, Stefan Reichelstein, Sebastian Schwenen, and Aleksandar Zaklan. “The impact of carbon disclosure mandates on emissions and financial operating performance.” *Review of Accounting Studies* 26, no.3 (2021): 1137-1175.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. “Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India.” *The Quarterly Journal of Economics* 128, no.4 (2013): 1499-1545.

Energy and Climate Intelligence Unit (ECIU). “Net Zero Scorecard.” (2024). Available at https://eciu.net/netzerotracker.

Environmental Protection Agency (EPA). “Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2021.” U.S. Environmental Protection Agency, EPA 430-R-23-002, 2023. https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks-1990-2021.

Ge, Weili and Sarah McVay. “The Disclosure of Material Weaknesses in Internal Control after the Sarbanes-Oxley Act.” *Accounting Horizons* 19, no.3 (2005): 137-158.

Gipper, Brandon, Samantha Ross, and Shawn X. Shi. “ESG Assurance in the United States.” Working Paper. Available at SSRN 4263085. (2023).

Glenk, Gunther. “Toward Decision-Useful Carbon Information.” Working Paper. Available at SSRN 4444037. (2023).

Goldhammer, Bernhard, Christian Busse, and Timo Busch. “Estimating Corporate Carbon Footprints with Externally Available Data.” *Journal of Industrial Ecology* 21, no.5 (2017): 1165-1179.

Griffin, Paul A., David H. Lont, and Estelle Y. Sun. “The Relevance to Investors of Greenhouse Gas Emission Disclosures.” *Contemporary Accounting Research* 34, no.2 (2017): 1265-1297.

Heitzman, Shane, and Mengjie Huang. “Internal Information Quality and the Sensitivity of Investment to Market Prices and Accounting Profits.” *Contemporary Accounting Research* 36, no.3 (2019): 1699-1723.

Jayaraman, Sudarshan and Joanna Shuang Wu. “Is Silence Golden? Real Effects of Mandatory Disclosure.” *The Review of Financial Studies* 32, no.6 (2019): 2225-2259.

Jones, Jennifer J. “Earnings Management During Import Relief Investigations.” *Journal of Accounting Research* 29, no.2 (1991): 193-228.

Kaplan, Robert S., and Karthik Ramanna. “Accounting for Climate Change.” *Harvard Business Review* 99, no.6 (2021): 120-131.

Kaplan, Robert S., and Karthik Ramanna. “How to Fix ESG Reporting.” Working Paper. Available at SSRN 3900146. (2021).

Klaaben, Lena and Christian Stoll. “Harmonizing corporate carbon footprints.” *Nature Communications* 12, no.1 (2021): 6149.

Kothari, S.P., Andrew J. Leone, and Charles E. Wasley. “Performance matched discretionary accrual measures.” *Journal of Accounting and Economics* 39, no.1 (2005): 163-197.

Lopez-de-Silanes, Florencio, Joseph A. McCahery, and Paul C. Pudschedl. "ESG Performance and Disclosure: A Cross-Country Analysis." Working Paper. Available at SSRN 3506084. (2019).

Luo, Le, Qingliang Tang, Hanlu Fan, and Jamie Ayers. "Corporate carbon assurance and the quality of carbon disclosure." *Accounting & Finance* 63, no.1 (2023): 657-690.

Maroun, Warren. "Does external assurance contribute to higher quality integrated reports?" *Journal of Accounting and Public Policy* 38, no. 4 (2019): 106670.

McNichols, Maureen F. "The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors: Discussion." *The Accounting Review* 77, (2002): 61-69.

Michelon, Giovanna, Dennis M. Patten, and Andrea M. Romi. "Creating legitimacy for sustainability assurance practices: Evidence from sustainability restatements." *European Accounting Review* 28, no. 2 (2019): 395-422.

Minnis, Michael. "The Value of Financial Statement Verification in Debt Financing: Evidence from Private U.S. Firms." *Journal of Accounting Research* 49, no.2 (2011): 457-506.

MSCI. "Filling the Blanks: Comparing Carbon Estimates Against Disclosures." MSCI ESG Research Issue Brief. (2016).

Muller, Nicholas Z. "Measuring Firm Environmental Performance to Inform ESG Investing." National Bureau of Economic Research Working Paper No. 29454. (2021).

Nguyen, Quyen, Ivan Diaz-Rainey, and Duminda Kuruppuarachchi. "Predicting corporate carbon footprints for climate finance risk analyses: A machine learning approach." *Energy Economics* 95 (2021).

Plumlee, Marlene, Darrell Brown, Rachel M. Hayes, and R. Scott Marshall. "Voluntary environmental disclosure quality and firm value: Further evidence." *Journal of Accounting and Public Policy* 34, no. 4 (2015): 336-361.

Refinitiv. "ESG Carbon Data and Estimate Models." (2017).

Rice, Sarah C. and David P. Weber. "How Effective Is Internal Control Reporting under SOX 404? Determinants of the (Non-)Disclosure of Existing Material Weaknesses." *Journal of Accounting Research* 50, no.3 (2012): 811-843.

Reichelstein, Stefan. "Carbon Emission Statements: Balance Sheets and Flow Statements." *Review of Accounting Studies*, forthcoming, 2023.

Schoenfeld, Jordan. "Cyber risk and voluntary Service Organization Control (SOC) audits." *Review of Accounting Studies*, forthcoming, 2023.

Watts, Ross L., and Jerold L. Zimmerman. "Agency problems, auditing, and the theory of the firm: Some evidence." *The Journal of Law and Economics* 26, no. 3 (1983): 613-633.

Whitaker, Michael, Garvin A. Heath, Patrick O'Donoughue, and Martin Vorum. "Life cycle greenhouse gas emissions of coal-fired electricity generation: Systematic review and harmonization." *Journal of Industrial Econology* 16 (2012): S53-S72.

Zhu, Christina. "Big data as a governance mechanism." *The Review of Financial Studies* 32, no.5 (2019): 2021-2061.

Appendix A. Variable Definitions

This table presents variable definitions for our empirical tests.

Variable	Definition
Abnormal Emissions	Absolute value of the residual term from a regression of the natural logarithm of scope-level emissions (from Trucost) on its determinants. See Section 3.1 for methodological details.
Assurance	A scope-specific indicator variable equal to one if either GRS data or CDP indicates that a firm obtains assurance over emissions for a particular emissions scope in a year, and zero otherwise.
Bloomberg E Disclosure Score	A proprietary score from Bloomberg based on the extent of a company's environmental disclosure. The score ranges from 0.1 (minimum amount of environmental disclosure) to 100 (disclosure of every data point collected by Bloomberg related to the environmental component of ESG). We divide scores by 100 in regression analyses to facilitate comparisons in our models.
Book-to-Market	The ratio of the book value of equity to the market value of equity, measured at the fiscal year-end, from Compustat.
Abnormal Accruals	A measure of firm abnormal accruals per McNichols (2002).
Falsification	An indicator variable equal to one if a country is headquartered in one of the eleven countries that require auditors to check the consistency of non-financial information with the financial statements, and zero otherwise.
Filing Lag	The natural logarithm of the difference between the date a firm submits its response to the CDP questionnaire and that year's CDP questionnaire release date.
Issues Identified	A scope-specific variable that counts the number of issues management identified with the carbon reporting system used to generate emissions.
Leverage	The ratio of total liabilities to total equity, measured at the fiscal year-end, from Compustat.
Loss	An indicator variable equal to one when basic earnings per share excluding extraordinary items from Compustat is less than zero, and zero otherwise.
Manager Uncertainty	A scope-specific decile rank of management's estimate of the level of uncertainty of the firm's total emissions. The question appeared in some years' CDP questionnaires, usually in Section 8 under "Data Accuracy."
Mandatory	An indicator variable equal to one if a country is headquartered in France, Italy, or Spain, and zero otherwise.
Meet or Beat	An indicator variable equal to one if income before extraordinary items is between 0 and 0.5% of total assets, and zero otherwise.
Omissions	A scope-specific variable that counts the number of sources of emissions which are not included in the number reported to the CDP. This question appeared in some years' CDP questionnaires, usually in Section 8 under "Scope 1 and 2 Emissions Data."
Post	An indicator variable equal to one after the implementation of NFRD in 2018, and zero otherwise.
Post Assurance	A scope-specific indicator variable equal to one for all years since the first instance of assurance, and zero otherwise.
Revisions	A scope-specific indicator variable equal to one if a firm revised its historical emissions number(s), and zero otherwise.
Refinitiv Disclosure Score	A proprietary score from Refinitiv of a company's ESG disclosures. The score ranges from 0 to 100 and is determined by Refinitiv's relative scoring algorithm. We divide scores by 100 in regression analyses to facilitate comparisons in our models.
ROA	Earnings before extraordinary items divided by average total assets, from Compustat.
Size	The natural logarithm of the total assets measured at fiscal year-end, from Compustat.

Appendix B. Examples of Issues in the Carbon Accounting System

This appendix provides examples of firm-provided anecdotes from CDP responses regarding issues in firms' carbon accounting systems. Responding firms select the categories.

Data Management

"Data management system is a combination of automated and some manual, Excel sheet-based calculations. Hence typographical errors, data handling errors are unavoidable, though we have ensured minimum error by our multi-level assurance check."

"We are currently using Excel data sheets to collect our data and are in the process of integrating our GHG emissions data into our financial data reporting system."

"The possibility of a human error occurring when entering data or using a calculation formula cannot be eliminated, because calculations are done using Excel instead of a calculation tool."

"Manual errors in reading meters, reading from invoices, transferring data onto Excel spreadsheets etc. although these are reviewed at a number of levels."

"Errors could occur when transferring data from supplier invoices, extrapolation of data within calculation models, and rounding errors."

Data Gaps

"Because we are a tenant in multi-tenant buildings we do not have direct control of or access to energy usage data for our facilities and, especially within the U.S. where we have the greatest concentration of facilities, our spaces are not separately submetered for utilities."

"As we are a global business and market-based reporting is a new level of detail, many electricity providers in many countries are not familiar with conversion factors or what they mean. It is therefore difficult to get these. Where they are unavailable, we follow the GRI Reporting hierarchy for market-based factors, meaning that the reported figures may not be as accurate as they should as a result of the data gaps."

Assumptions and Extrapolations

"Data gathered was for entire buildings and then Capital Power's portion was determined based on square footage."

"Where we do not have direct access to the data we rely on the building landlord to provide total building energy usage for the building, which we then prorate for our applicable portion of the total building space. When we are not able to obtain data from a landlord we must estimate energy usage using published energy intensity factors appropriate for each region."

"This location shares utilities with a separate company. We split utilities by floor space. We claim 67% of the total utility supplied to the facility."

"71% of scope 1 and 2 emissions come from the energy bought for real estate investment activities and for corporate use. These emissions are calculated based on invoices collected for all the reporting scope. The rest of scope 1 and 2 emissions come from business travel done by employees. These emissions are calculated based on invoices collected or on estimated kilometers done. This can lead to higher uncertainty compared to data read on invoices."

Sampling

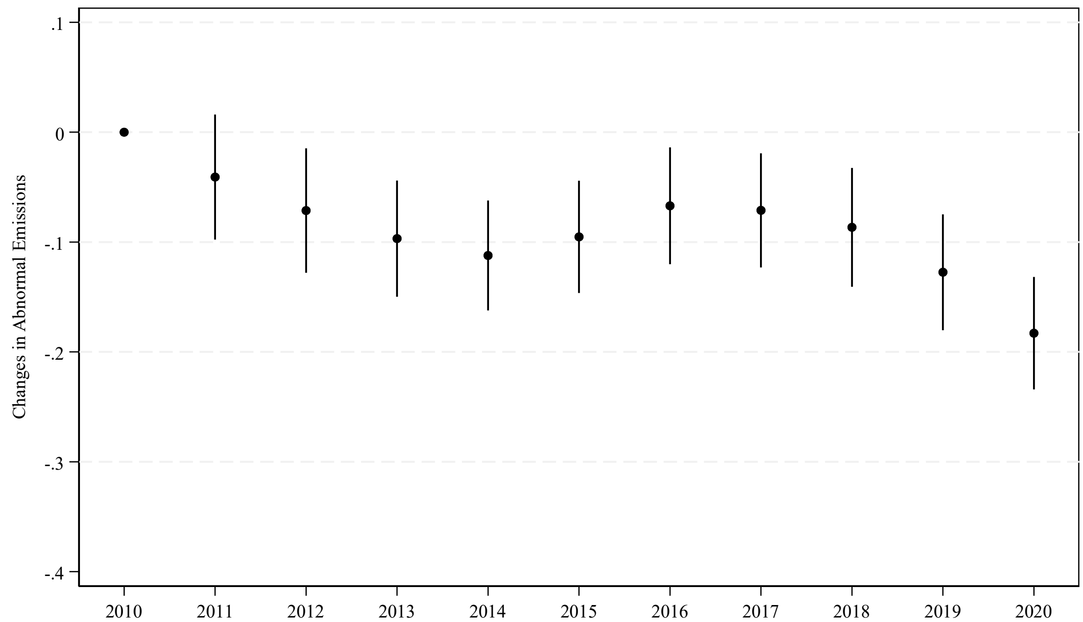
"Calculation based on electric bill for a sample 3 month period."

"To calculate the standard CO_2 emissions produced at the company's work sites, we took a sample that was equivalent to 20%–30% of the completed construction amount. We calculate the total CO_2 emissions amount based on the standard CO_2 emissions calculated based on that sample."

Figure 1. Intertemporal Changes in Abnormal Emissions

This figure presents time trends of carbon accounting quality. For each scope, *Abnormal Emissions* is regressed on calendar-year indicators, with 2010 omitted as the base year. Coefficient estimates for year indicators (dots) and 95% confidence intervals (bars) are presented for each coefficient. Panels A and B report results for Scopes 1 and 2, respectively. All calculations are based on the full sample of data at the firm-year level.

Panel A. Scope 1



Panel B. Scope 2

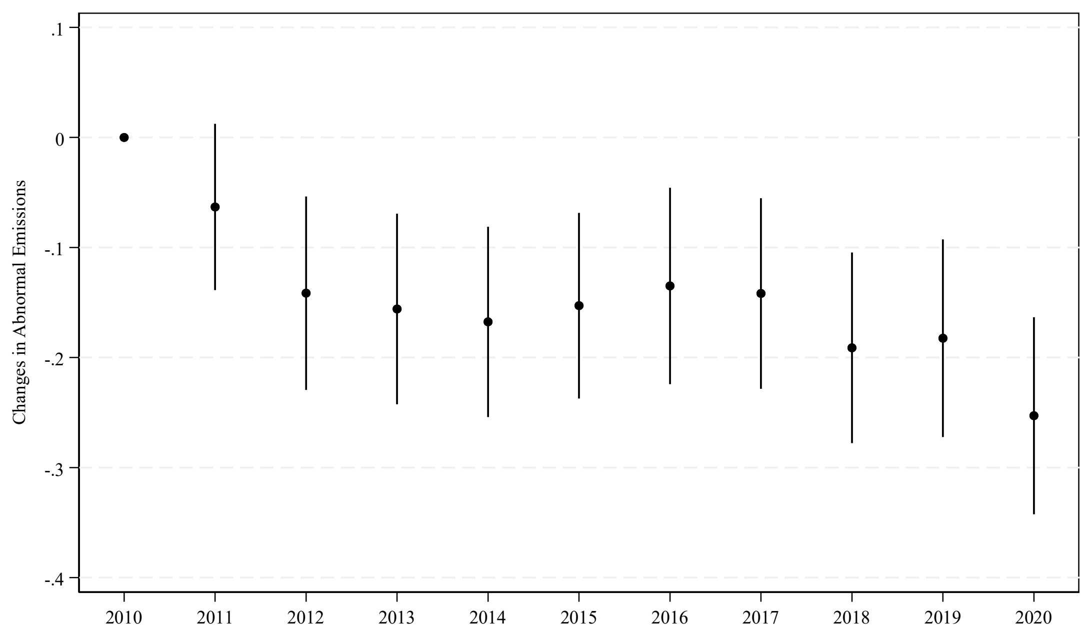
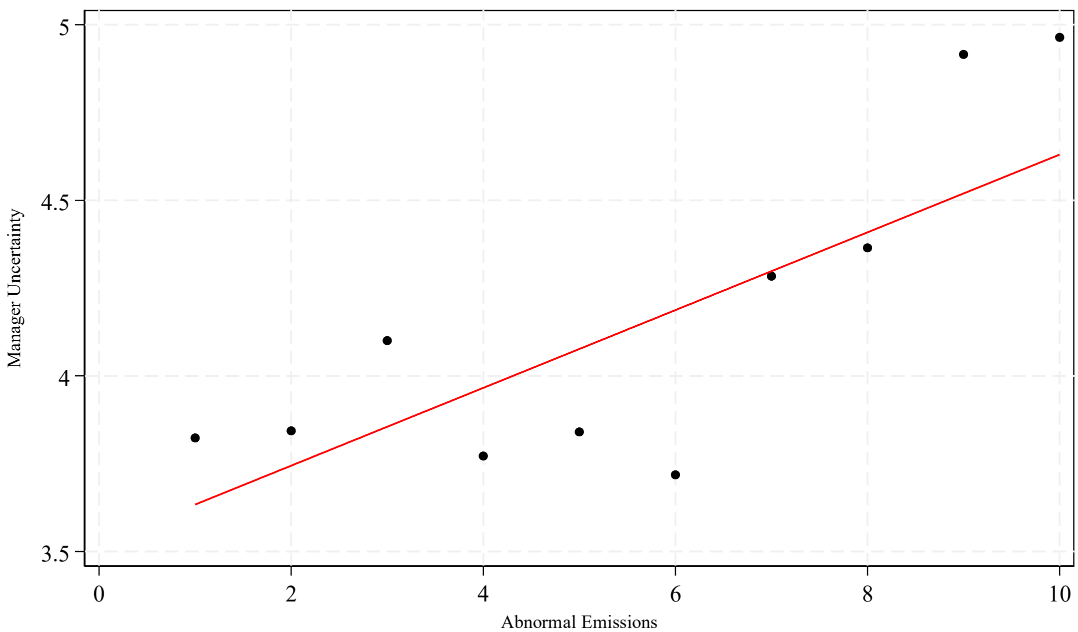


Figure 2. Measure Validation

This figure presents binned scatterplots of *Abnormal Emissions* and *Manager Uncertainty*. Panels A and B report results for Scopes 1 and 2, respectively. To generate binned scatterplots, *Abnormal Emissions* is sorted into ten equal-sized bins. Within each bin, we calculate and plot the mean of *Manager Uncertainty*. The red line is the line of best fit generated using underlying data.

Panel A. Scope 1



Panel B. Scope 2

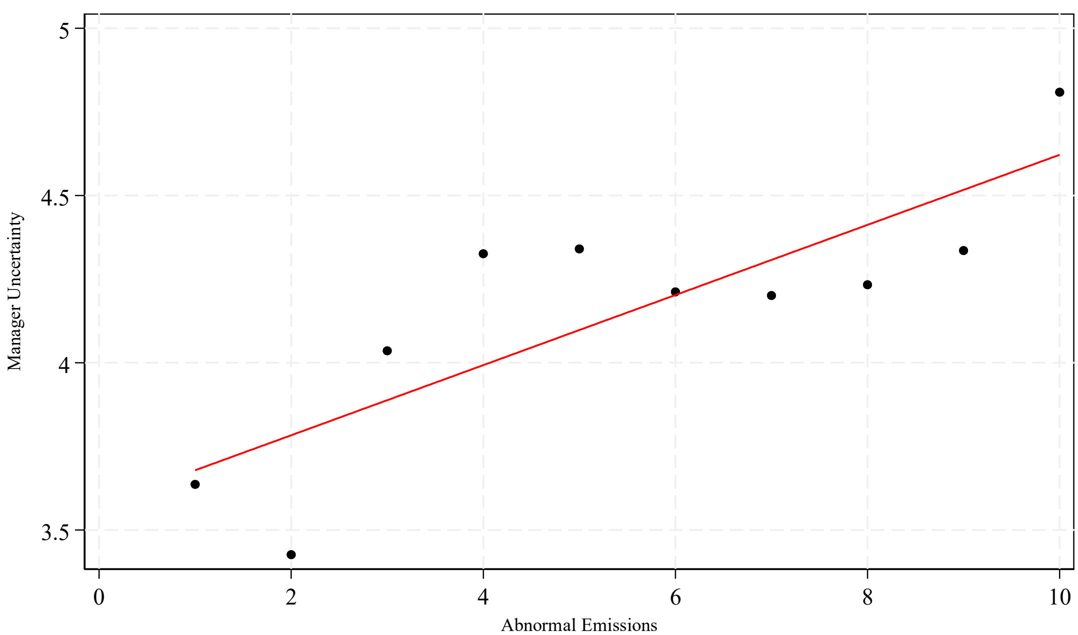
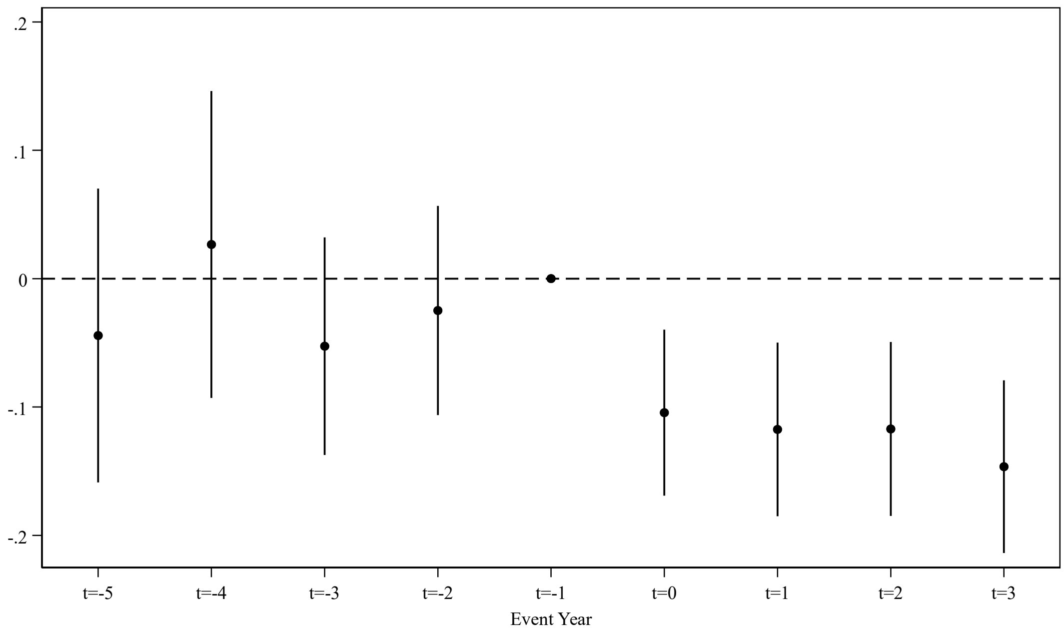


Figure 3. Effect of First-Time Assurance on Abnormal Emissions in Event Time

This figure presents event-time analysis of the effect of first-time scope-specific assurance on abnormal emissions. The first year with assurance is t=0. All coefficients are benchmarked to the year prior to assurance adoption (t=-1). Coefficient estimates (dots) and 95% confidence intervals (bars) are presented for each coefficient.

Panel A. Scope 1



Panel B. Scope 2

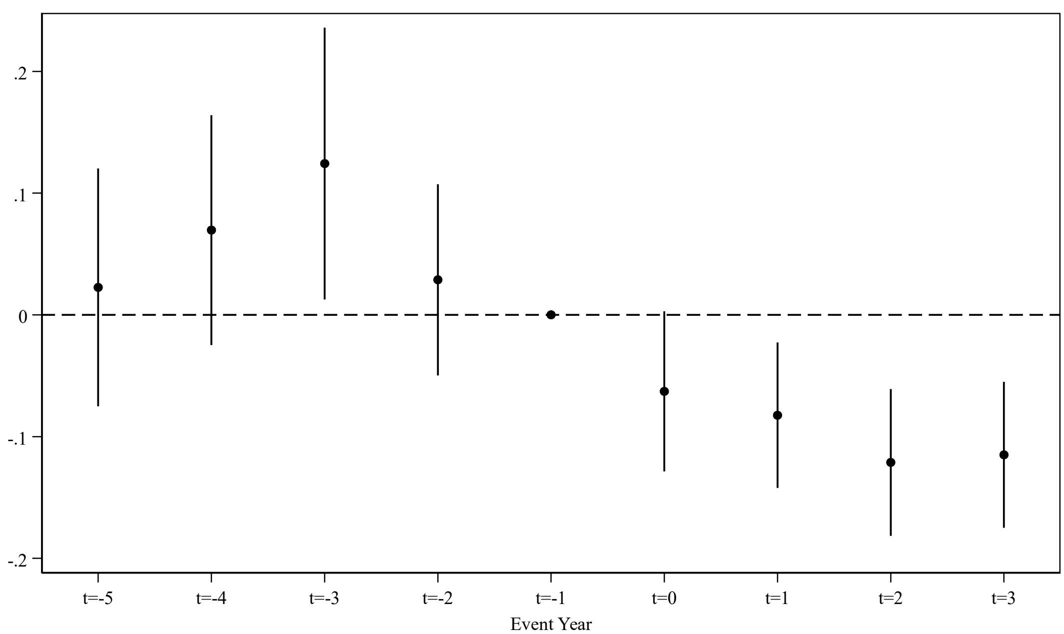


Table 1. Descriptive Statistics

This table presents descriptive statistics of key variables used in the study. All continuous variables are winsorized at the 1st and 99th percentiles.

		(1) N	(2) Mean	(3) SD	(4) 25th	(5) Median	(6) 75th
Firm Characteristics							
Size		5,395	9.83	1.48	8.79	9.77	10.87
ROA		5,395	0.04	0.09	0.01	0.04	0.08
Book-to-Market		5,395	0.57	0.56	0.23	0.42	0.73
Leverage		5,395	3.04	5.66	0.99	1.71	3.20
Loss		5,395	0.15	0.35	0	0	0
Meet or Beat		3,821	0.04	0.19	0	0	0
Abnormal Accruals		3,003	9.03	10.57	1.91	5.20	12.63
Carbon Emissions Reporting and Assurance							
Assurance	Scope 1	5,395	0.49	0.50	0	0	1
	Scope 2	5,202	0.48	0.50	0	0	1
Abnormal Emissions	Scope 1	5,395	0.32	0.45	0.08	0.18	0.36
	Scope 2	5,202	0.31	0.47	0.08	0.17	0.34
CDP Responses							
Issues Identified	Scope 1	3,517	1.99	1.31	1.00	2.00	3.00
	Scope 2	3,454	1.91	1.31	1.00	2.00	3.00
Omissions	Scope 1	5,215	0.45	0.50	0	0	1
	Scope 2	5,215	0.37	0.48	0	0	1
Revisions	Scope 1	5,215	0.05	0.22	0	0	0
	Scope 2	5,215	0.03	0.18	0	0	0
Manager Uncertainty	Scope 1	3,305	4.28	2.89	1	4	7
	Scope 2	3,184	4.27	2.90	1	4	7
Filing Lag		4,733	4.88	0.24	4.75	4.80	5.11
Scores from ESG Rating Agencies							
Refinitiv Disclosure Score		3,471	73.12	27.98	79.29	82.57	86.24
Bloomberg E Disclosure Score		1,686	34.17	15.67	23.26	34.88	46.51

Table 2. Measure Validation

This table presents results from regression analyses of model-estimated carbon accounting quality (*Abnormal Emissions*) on firm perception of carbon accounting quality (*Manager Uncertainty*) or disclosure scores from ESG rating agencies. Definitions of all variables are as described in Appendix A. t–statistics appear in parentheses and are clustered by firm. *, **, *** indicate statistical significance (two–sided) at the 0.1, 0.05, and 0.01 levels, respectively.

Panel A. Manager Uncertainty

	(1)	(2)
	Dependent Variable: *Abnormal Emissions*	
Manager Uncertainty	0.005***	0.004***
	(3.96)	(3.59)
Sample	Scope 1	Scope 2
Observations	2,064	1,940
R^2	0.012	0.009

Panel B. Disclosure Scores from ESG Rating Agencies

	(1)	(2)	(3)	(4)
	Dependent Variable: *Abnormal Emissions*			
Refinitiv Disclosure Score	-0.066***	-0.036***	-	-
	(-5.38)	(-3.31)		
Bloomberg E Disclosure Score	-	-	-0.147***	-0.078***
			(-4.93)	(-2.63)
Sample	Scope 1	Scope 2	Scope 1	Scope 2
Observations	3,471	3,315	1,686	1,604
R^2	0.013	0.005	0.022	0.008

Table 3. Assurance and Carbon Accounting Quality

This table presents results from regressions of carbon accounting quality measures on assurance. Panel A reports the main result. Panel B reports results using two alternative measures of carbon accounting quality, firm perception of reporting uncertainty in Scopes 1 and 2 emissions, and the number of days it takes for a firm to submit its questionnaire response to the CDP. *Abnormal Emissions, Manager Uncertainty* and *Assurance* are all scope-specific, with the scope used indicated at the bottom of the table. *Filing Lag* varies only by firm-year pairs. Definitions of all variables are as described in Appendix A. t–statistics appear in parentheses and are clustered by firm. *, **, *** indicate statistical significance (two-sided) at the 0.1, 0.05, and 0.01 levels, respectively.

Panel A. Main Result

	(1)	(2)	(3)	(4)
	Dependent Variable: *Abnormal Emissions*			
Assurance	-0.162***	-0.082***	-0.163***	-0.104***
	(-11.39)	(-5.27)	(-10.47)	(-6.57)
Controls:				
Size	-	-0.067***	-	-0.031***
		(-9.98)		(-4.73)
ROA	-	-0.055	-	-0.117
		(-0.51)		(-0.92)
Book-to-Market	-	-0.011	-	0.038*
		(-0.82)		(1.96)
Leverage	-	0.002	-	-0.001
		(1.45)		(-0.87)
Loss	-	0.026	-	0.027
		(0.95)		(0.88)
Sample	Scope 1	Scope 1	Scope 2	Scope 2
Fixed Effects	No	Ind, Year	No	Ind, Year
Observations	5,395	5,395	5,202	5,202
R^2	0.033	0.096	0.030	0.145

Panel B. Alternative Measures of Carbon Accounting Quality

	(1)	(2)	(3)	(4)
	Dependent Variable:			
	Manager Uncertainty		*Filing Lag*	
Assurance	-0.401**	-0.687***	-0.011*	-0.009*
	(-2.14)	(-3.71)	(-1.93)	(-1.69)
Sample	Scope 1	Scope 2	Scope 1	Scope 2
Controls	Yes	Yes	Yes	Yes
Ind, Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	3,305	3,184	4,733	4,733
R^2	0.068	0.090	0.712	0.712

Table 4. Addressing Alternative Explanations

This table presents results from analyses addressing potential alternative explanations. Panel A reports results using alternative fixed effects and research designs. Panel B reports results from falsification tests. Definitions of all variables are as described in Appendix A. t–statistics appear in parentheses and are clustered by firm. *, **, *** indicate statistical significance (two-sided) at the 0.1, 0.05, and 0.01 levels, respectively.

Panel A. Alternative Research Designs

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: *Abnormal Emissions*					
Assurance	-0.032*	-	-	-0.049***	-	-
	(-1.94)			(-2.91)		
Post Assurance	-	-0.140***	-0.065**	-	-0.148***	-0.067**
		(-8.31)	(-2.39)		(-7.94)	(-2.25)
Sample	Scope 1	Scope 1	Scope 1	Scope 2	Scope 2	Scope 2
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Firm, Year	Ind, Year	Firm, Year	Firm, Year	Ind, Year	Firm, Year
Observations	5,205	5,395	5,205	5,017	5,202	5,017
R^2	0.367	0.108	0.368	0.421	0.155	0.421

Panel B. Falsification Tests

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable:					
	Abnormal Emissions		*Meet or Beat*		*Abnormal Accruals*	
Scope 1 Assurance	-0.072**	0.020	-0.003	-	-0.035	-
	(-2.44)	(0.30)	(-0.47)		(-0.07)	
Scope 2 Assurance	-0.011	-0.123*	-	-0.008	-	0.525
	(-0.37)	(-1.82)		(-1.24)		(1.08)
Sample	Scope 1	Scope 2	Scope 1	Scope 2	Scope 1	Scope 2
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Ind, Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,395	5,202	3,821	3,682	3,003	2,884
R^2	0.096	0.145	0.171	0.177	0.191	0.185

Table 5. Cross-sectional Variation in Assurance Quality

This table presents cross-sectional analyses exploring whether the relation between assurance and carbon accounting quality varies by assurance quality. [1] Refers to a limited level of scope-level assurance (Scope 1 in column (5) and Scope 2 in column (6)) and assurance is not provided over all three scopes. [2] Refers to a limited level of scope-level assurance and assurance is provided over all three scopes. [3] Refers to a reasonable level of scope-level assurance and assurance is not provided over all three scopes. [4] Refers to a reasonable level of scope-level assurance and assurance is provided over all three scopes. Definitions of all variables are as described in Appendix A. t–statistics appear in parentheses and are clustered by firm. *, **, *** indicate statistical significance (two-sided) at the 0.1, 0.05, and 0.01 levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: *Abnormal Emissions*					
[A] *Limited Assurance*	-0.080***	-0.101***	-	-	-	-
	(-4.89)	(-6.27)				
[B] *Reasonable Assurance*	-0.090***	-0.117***	-	-	-	-
	(-4.63)	(-5.32)				
[A] *Some Scopes Assured*	-	-	-0.076***	-0.088***	-	-
			(-4.04)	(-4.60)		
[B] *All Scopes Assured*	-	-	-0.085***	-0.110***	-	-
			(-5.06)	(-6.56)		
[1] *Limited + Some*	-	-	-	-	-0.072***	-0.088***
					(-3.60)	(-4.24)
[2] *Limited + All*	-	-	-	-	-0.083***	-0.087***
					(-3.01)	(-3.18)
[3] *Reasonable + Some*	-	-	-	-	-0.083***	-0.105***
					(-4.68)	(-6.23)
[4] *Reasonable + All*	-	-	-	-	-0.095***	-0.131***
					(-4.38)	(-5.10)
Pr([B]>[A]) p-value:	0.262	0.179	0.281	0.083		
Pr([4]>[1]) p-value:					0.162	0.053
Sample	Scope 1	Scope 2	Scope 1	Scope 2	Scope 1	Scope 2
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Ind, Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,395	5,202	5,395	5,202	5,395	5,202
R^2	0.096	0.145	0.096	0.145	0.096	0.146

Table 6. Mechanism: Assurance and Indicators of the Carbon Accounting System

This table presents results from an analysis of the association between assurance and three indicators of the carbon accounting system: the number of issues identified (*Issues Identified*), the probability that at least one emissions source is omitted (*Omissions*), and whether historical GHG emissions numbers are updated (*Revisions*). Definitions of all variables are as described in Appendix A. t–statistics appear in parentheses and are clustered by firm. *, **, *** indicate statistical significance (two-sided) at the 0.1, 0.05, and 0.01 levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable:					
	Issues Identified		*Omissions*		*Revisions*	
Assurance	0.248***	0.291***	-0.059**	-0.119***	0.015*	0.020***
	(3.01)	(3.59)	(-2.21)	(-4.56)	(1.88)	(3.13)
Sample	Scope 1	Scope 2	Scope 1	Scope 2	Scope 1	Scope 2
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Ind, Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,517	3,454	5,215	5,215	5,215	5,215
R^2	0.080	0.067	0.040	0.046	0.053	0.035

Table 7. International Evidence

This table presents results from an analysis of the relation between assurance and carbon accounting quality using a sample of non-U.S. firms. Panel A reports sample composition. Panel B reports descriptive statistics. Panel C reports regressions of abnormal emissions on assurance. Definitions of all variables are as described in Appendix A. t–statistics appear in parentheses and are clustered by firm. *, **, *** indicate statistical significance (two-sided) at the 0.1, 0.05, and 0.01 levels, respectively.

Panel A. Sample

	Scope 1		Scope 2	
Country	Observations	Percent	Observations	Percent
Australia	369	3.47%	364	3.44%
Austria	106	1.00%	92	0.87%
Belgium	106	1.00%	101	0.96%
Brazil	182	1.71%	174	1.65%
China	445	4.19%	561	5.31%
Croatia	11	0.10%	8	0.08%
Cyprus	3	0.03%	2	0.02%
Czech Republic	10	0.09%	3	0.03%
Denmark	165	1.55%	155	1.47%
Estonia	4	0.04%	5	0.05%
Finland	219	2.06%	207	1.96%
France	701	6.59%	701	6.63%
Germany	507	4.77%	489	4.63%
Greece	56	0.53%	40	0.38%
Hong Kong	511	4.81%	615	5.82%
Hungary	19	0.18%	17	0.16%
India	281	2.64%	277	2.62%
Indonesia	10	0.09%	10	0.09%
Ireland	98	0.92%	91	0.86%
Italy	283	2.66%	271	2.56%
Japan	2,264	21.29%	2,195	20.77%
Luxembourg	67	0.63%	66	0.62%
Malaysia	70	0.66%	83	0.79%
Mexico	57	0.54%	61	0.58%
Netherlands	215	2.02%	206	1.95%
New Zealand	11	0.10%	11	0.10%
Norway	191	1.80%	173	1.64%
Philippines	19	0.18%	21	0.20%
Poland	52	0.49%	44	0.42%
Portugal	73	0.69%	55	0.52%
Romania	4	0.04%	2	0.02%
Russia	7	0.07%	2	0.02%
Singapore	32	0.30%	33	0.31%
Slovenia	6	0.06%	7	0.07%
South Africa	385	3.62%	381	3.61%
Spain	279	2.62%	274	2.59%
Sweden	376	3.54%	354	3.35%
Switzerland	303	2.85%	275	2.60%
Taiwan	126	1.19%	147	1.39%
Thailand	27	0.25%	26	0.25%
Turkey	85	0.80%	80	0.76%
United Kingdom	1,897	17.84%	1,889	17.87%
Total	10,632	100%	10,568	100%

continued

Table 7. International Evidence—continued

Panel B. Descriptive Statistics

		(1) N	(2) Mean	(3) SD	(4) 25^{th}	(5) Median	(6) 75^{th}
Firm Characteristics							
Size		10,632	9.93	2.68	7.95	9.53	11.96
ROA		10,632	0.04	0.08	0.02	0.04	0.07
Book-to-Market		10,632	0.89	1.43	0.32	0.58	0.98
Leverage		10,632	0.55	0.20	0.42	0.55	0.68
Loss		10,632	0.12	0.33	0	0	0
Carbon Emissions Reporting and Assurance							
Assurance	Scope 1	10,632	0.30	0.46	0	0	1
	Scope 2	10,568	0.28	0.45	0	0	1
Abnormal Emissions	Scope 1	10,632	0.37	0.54	0.09	0.20	0.40
	Scope 2	10,568	0.36	0.52	0.09	0.21	0.42

Panel C. Regression Analysis

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: *Abnormal Emissions*					
Assurance	-0.056***	-0.010	-	-0.038***	-0.003	-
	(-4.97)	(-0.60)		(-3.08)	(-0.17)	
Post Assurance	-	-	-0.068***	-	-	-0.049*
			(-2.97)			(-1.85)
Sample	Scope 1	Scope1	Scope 1	Scope 2	Scope 2	Scope 2
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Ind, Year, Country	Firm, Year	Firm, Year	Ind, Year, Country	Firm, Year	Firm, Year
Observations	10,632	10,219	10,219	10,568	10,133	10,133
R^2	0.120	0.397	0.398	0.094	0.398	0.399

Table 8. NFRD and Carbon Accounting Quality

This table presents results from analyses of the effect of NFRD on assurance and abnormal emissions. In columns (1) and (2), the sample is all European Union (E.U.) member countries for which we have data plus Norway and the United Kingdom. In columns (3) and (4), the sample is E.U. countries whose sustainability disclosures are subject to mandatory assurance (France, Italy, and Spain), as well as E.U. member countries (plus Norway and the U.K.) for which we have data and whose sustainability disclosures are subject to auditor consistency checks (Austria, Belgium, Denmark, Finland, France, Germany, Netherlands, Norway, Slovenia, Sweden, and the United Kingdom). In columns (5) and (6), the sample is all E.U. member countries (plus Norway and the U.K.) for which we have data, excluding France, Italy, and Spain. In all columns, we restrict our sample to data after 2016. Definitions of all variables are as described in Appendix A. t–statistics appear in parentheses and are clustered by firm. *, **, *** indicate statistical significance (two-sided) at the 0.1, 0.05, and 0.01 levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: *Abnormal Emissions*					
Mandatory × Post	-0.107***	-0.079*	-0.118***	-0.089**	-	-
	(-2.78)	(-1.83)	(-3.01)	(-2.07)		
Falsification × Post	-	-	-	-	0.020	0.138
					(0.48)	(1.04)
Sample	Scope 1	Scope 2	Scope 1	Scope 2	Scope 1	Scope 2
Firms	All E.U. Countries		E.U. Countries with Mandatory Assurance or Auditor Consistency Checks		Excluding E.U. Countries with Mandatory Assurance	
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm, Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,343	2,257	2,163	2,119	1,696	1,603
R^2	0.488	0.501	0.491	0.492	0.503	0.515

Internet Appendix

Carbon Accounting Quality: Measurement and the Role of Assurance

Brandon Gipper, Fiona Sequeira, Shawn X. Shi

Table of Contents

Table IA-1. Variable Definitions

This table presents definitions for new variables used in our Internet Appendix. Variables used in our IA that are also used in the main analyses are not defined below; please refer to Appendix A of the paper.

Variable	Definition
Asset Tangibility	Net property, plant, and equipment (PP&E) scaled by total assets, from Compustat.
Business Centrality	The natural logarithm of one plus the number of unique states in which the firm operates, from WRDS Segment data.
CapEx	Capital expenditures scaled by total assets, from Compustat.
Dividends	The total amount of dividends scaled by net income, from Compustat.
Energy Intensity	Emission intensity of the energy mix of the firm (g CO2/kWh).
EPS Growth	The year-over-year change in income before extraordinary items, from Compustat.
HHI or PCM	Total sales minus total cost of goods sold, divided by total sales, from Compustat.
Inst Own	The fraction of the firm's equity owned by institutional investors, from Compustat.
Intangibles	The total intangible assets measured at fiscal year-end, from Compustat.
Ln(*Assets*)	The natural logarithm of the total assets measured at fiscal year-end, from Compustat.
Ln(*Asset Age*)	The natural logarithm of one plus the difference between gross property, plant, and equipment (PP&E) and net PP&E divided by annual depreciation, from Compustat.
Ln(*COGS*)	The natural logarithm of one plus the total cost of goods sold, from Compustat.
Ln(*Emissions*)	The natural logarithm of total scope-level emissions, from Trucost.
Ln(*Employees*)	The natural logarithm of one plus the count of employees, from Compustat.
Ln(*Inventory*)	The natural logarithm of one plus the total amount of inventory, from Compustat.
Ln(*Lagged Emissions*)	The level of historical scope-level emissions; i.e. the total amount of Scope 1 or Scope 2 emissions from the prior year, from Trucost.
Ln(*Lagged Inventory*)	The natural logarithm of one plus the amount of total inventory from the prior year, from Compustat.
Ln(*MVE*)	The natural logarithm of the number of common shares outstanding multiplied by the share price, from Compustat.
Ln(*PPE*)	The natural logarithm of net property, plant, and equipment (PP&E), from Compustat.
Ln(*Sales*)	The natural logarithm of total sales, from Compustat.
Return	The firm's stock return over the year, from Compustat.
Sales Growth	The year-over-year change in firm sales, from Compustat.

Table IA-2. Corporate Carbon Emissions Models in Prior Studies

This table explores various models of corporate carbon emissions. Below, we present a survey of the determinants used by prior studies and data vendors to model carbon emissions, as well as R^2 from simple regressions of Scope 1 emissions on each determinant. GBB is short for Goldhammer et al. (2017). GLS is short for Griffin et al. (2017). DERSZ is short for Downar et al. (2021). BK is short for Bolton and Kacperczyk (2021). NDK is short for Nguyen et al. (2021). CKOR is short for Cohen et al. (2023a). CKO is short for Cohen et al. (2023b). All models below use linear methods, except for NDK (2021), which uses a machine learning approach. Individual R^2 for Energy Intensity is missing because the variable is defined at the country level. Definitions of all variables are as described in Table IA-1.

	Prior Studies and Data Vendors										
Determinants	GBB 2017	GLS 2017	DERSZ 2021	BK 2021	NDK 2021	CKOR 2023	CKO 2023	CDP	MSCI	Refinitiv	Individual R^2
Ln(Lagged Emissions)					×				×	×	0.955
Ln(PP&E)				×	×						0.474
Asset Tangibility	×		×		×	×	×				0.403
Ln(Asset Age)		×			×						0.283
CapEx		×		×	×						0.170
Ln(Inventory)											0.161
Ln(COGS)											0.134
Ln(Sales)	×	×		×	×	×		×	×	×	0.094
Ln(Assets)					×	×	×				0.059
Ln(MVE)			×	×							0.025
Inst Own						×					0.023
Ln(Employees)					×					×	0.018
Leverage		×	×	×		×	×				0.018
Business Centrality	×										0.013
Dividends						×	×				0.010
Return						×	×				0.010
Book-to-Market			×								0.010
Sales Growth				×							0.005
Intangibles		×			×						0.004
ROA				×		×	×				0.002
EPS Growth				×							0.001
HHI or PCM	×	×		×	×						0.001
Energy Intensity	×										N/A

Table IA-3. Corporate Carbon Emissions Model

This table presents results from regressions of corporate carbon emissions on the determinants used in this study. In our analysis in the manuscript, we allow the coefficient for each determinant to vary by industry. Below, we present results without industry-level coefficients for parsimony. Definitions of all variables are as described in Table IA-1. t–statistics appear in parentheses and are clustered by firm. *, **, *** indicate statistical significance (two-sided) at the 0.1, 0.05, and 0.01 levels, respectively.

	(1)	(2)	(3)	(4)
	Dependent Variable: Ln(*Emissions*)			
	U.S. Firms		Non-U.S. Firms	
Historical Emissions				
Ln(*Lagged Emissions*)	0.913***	0.856***	0.916***	0.848***
	(99.92)	(41.82)	(123.01)	(76.35)
Production Function				
Ln(*PPE*)	0.093***	0.038*	0.079***	0.118***
	(3.91)	(1.72)	(3.25)	(4.18)
Asset Tangibility	0.427***	0.495***	0.247**	0.028
	(3.66)	(3.58)	(2.35)	(0.24)
Ln(*Asset Age*)	-0.005	-0.019	0.034**	0.053***
	(-0.27)	(-0.84)	(2.22)	(3.19)
CapEx	0.260	-0.127	0.316	0.720**
	(0.78)	(-0.27)	(1.13)	(2.23)
Ln(*Size*)	-0.001	-0.008	0.001	-0.017
	(-0.03)	(-0.33)	(0.04)	(-0.56)
Ln(*Employees*)	-0.016	0.078***	0.098***	0.153***
	(-0.93)	(3.36)	(7.64)	(9.69)
Business Centrality	0.007	0.013		
	(0.87)	(1.36)		
Production Output				
Ln(*Inventory*)	0.033*	0.040**	0.025	0.041
	(1.91)	(2.34)	(0.91)	(1.50)
Ln(*Lagged Inventory*)	-0.024	-0.029*	-0.015	-0.028
	(-1.40)	(-1.78)	(-0.58)	(-1.06)
Ln(*COGS*)	0.029	-0.010	0.066***	0.018
	(1.25)	(-0.44)	(3.99)	(1.14)
Ln(*Sales*)	0.001	0.048	-0.112***	-0.078***
	(0.04)	(1.17)	(-3.96)	(-2.81)
Sample	Scope 1	Scope 2	Scope 1	Scope 2
Ind × Year Fixed Effects	Yes	Yes	No	No
Country × Ind × Year Fixed Effects	No	No	Yes	Yes
Observations	5,414	5,220	10,640	10,576
R^2	0.959	0.893	0.950	0.921

Table IA-4. Main Analysis Robustness

This table presents results from an analysis exploring the robustness of the main finding. Panel A presents various modifications of our approach using OLS regressions, and Panel B presents an alternative design using quantile regressions. In Panel A, we present seven modifications. Modifications [1] to [4] are related to different ways to model corporate carbon emissions in Eq. (1). In modification [1], instead of allowing each determinant to have separate coefficients for each industry, we allow each determinant to have separate coefficients for each year in our sample. In modification [2], we allow each determinant to have separate coefficients for each industry and half-decade pairs. In modification [3], we do not allow determinants to have separate coefficients for each industry. In modification [4], we add scope-level assurance as an additional determinant in our main emissions model. In modification [5], we use changes in emissions, instead of emissions levels, as the dependent variable in the corporate carbon emissions model. Accordingly, we drop lagged emissions as a determinant. In modification [6], we run our Table 3 regression analyses with industry-by-emissions model regressors added as additional controls as well as industry-by-year fixed effects, both using the same industry definition that we used in the emissions model. In modification [7], we perform a principal component analysis of our three carbon accounting quality measures, i.e., *Abnormal Emissions*, *Manager Uncertainty*, and *Filing Lag*, to create a composite carbon accounting quality (CAQ) measure from the first principal component (PC). In columns (1) and (2), we report the coefficient on *Assurance* from estimating Eq. (2) for Scopes 1 and 2, respectively. Panel B presents results from a single-step quantile regression that includes all of the emissions model and Table 3 regressors from our main specification, an approach proposed by Chen et al. (2023) to avoid the potential bias associated with the use of absolute residuals as dependent variables. With this approach, we estimate the relation between the model regressors and the dependent variable (Ln(*Emissions*)) at various points of the conditional distribution of scope-level emissions; namely, at the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles. Definitions of all variables are as described in Appendix A and Table IA-1. t–statistics appear in parentheses and are clustered by firm. *, **, *** indicate statistical significance (two-sided) at the 0.1, 0.05, and 0.01 levels, respectively.

Panel A. Model Modifications

Modification:	(1)	(2)
	Coefficient on *Assurance*	
[1] Separate coefficients for each year in emissions model (instead of separate coefficients for each industry)	-0.089***	-0.113***
	(-5.69)	(-6.52)
[2] Separate coefficients for each industry and half-decade pair in emissions model (instead of separate coefficients for each industry)	-0.080***	-0.105***
	(-5.27)	(-6.80)
[3] No separate coefficients for each industry in emissions model	-0.086***	-0.116***
	(-5.36)	(-6.40)
[4] Scope-level assurance added as emissions model determinant	-0.084***	-0.104***
	(-5.40)	(-6.55)
[5] Changes in emissions as emissions model dependent variable	-0.092***	-0.129***
	(-5.59)	(-7.18)
[6] All determinants of carbon emissions included as controls	-0.047***	-0.071***
	(-3.16)	(-4.70)
[7] First PC of the three CAQ measures as dependent variable	-0.179**	-0.241***
	(-2.34)	(-3.00)
Sample	Scope 1	Scope 2
Controls	Yes	Yes
Ind, Year Fixed Effects	Yes	Yes

Table IA-4. Main Analysis Robustness—continued

Panel B. Quantile Regressions

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: Ln(*Emissions*)				
	Q10	Q25	Q50	Q75	Q90
Scope 1:					
Assurance	0.094**	0.037	0.000	-0.031*	-0.068***
	(2.47)	(1.58)	(-0.03)	(-1.68)	(-2.83)
Scope 2:					
Assurance	0.063**	0.002	-0.051***	-0.106***	-0.187***
	(2.13)	(0.08)	(-2.63)	(-4.58)	(-5.38)

Table IA-5. Additional Cross-sectional Analyses

This table presents cross-sectional analyses exploring whether the relation between assurance and carbon accounting quality varies by assuror identity and the presence of emissions reduction targets. Definitions of all variables are as described in Appendix A and Table IA-1. t–statistics appear in parentheses and are clustered by firm. *, **, *** indicate statistical significance (two-sided) at the 0.1, 0.05, and 0.01 levels, respectively.

	(1)	(2)	(3)	(4)
	Dependent Variable: *Abnormal Emissions*			
Split Variables:	Assurance Provided by Financial Auditor?		Firm has Emissions Reduction Target(s)?	
No	-0.066***	-0.081***	-0.067***	-0.096***
	(-3.61)	(-3.38)	(-3.95)	(-5.39)
Yes	-0.061*	-0.057	-0.102***	-0.115***
	(-1.96)	(-1.59)	(-5.88)	(-7.21)
Pr(*Yes*=*No*) p-value:	0.857	0.428	0.014	0.145
Sample	Scope 1	Scope 2	Scope 1	Scope 2
Controls	Yes	Yes	Yes	Yes
Ind, Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	2,282	2,169	5,395	5,202
R^2	0.059	0.219	0.097	0.145