Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

How third-party assurance changes corporate emissions reporting and firm behavior

Fernando Loaiza

fernando.loaiza@ip.mpg.de

Max Planck Institute for Innovation and Competition https://orcid.org/0000-0002-1428-6682

Benedict Probst

Max Planck Institute for Innovation and Competition https://orcid.org/0000-0002-1149-8938

Article

Keywords:

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11479006/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

How third-party assurance changes corporate emissions reporting and firm behavior

Abstract

Mandatory third-party assurance of corporate emissions is spreading, but it is unclear whether verification changes emissions themselves or only the figures firms report. We estimate the causal effect of third-party verification of corporate emissions using staggered adoption among CDP-reporting firms from 2018 to 2025. Verification produces two distinct effects on how firms report and how they operate. A reporting effect appears immediately in the extensive and intensive margins of Scope 3 value-chain categories, changing category coverage and reported emissions. A behavioral effect operates through Scope 1, as emissions fall gradually in the years after adoption, driven by a shift away from fuel consumption. Scope 2 emissions, both location-based and market-based, show no robust change. These findings suggest that third-party verification can improve emissions inventories, influence subsequent operational decisions and guide the design of mandatory assurance requirements for corporate climate reporting.

Third-party assurance of corporate climate reporting is expanding rapidly. Even before recent mandates, third-party assurance had developed into an established international practice in sustainability reporting [1], and has become increasingly common both across ESG reporting among large US firms and specifically for corporate carbon-emissions disclosures internationally [2, 3]. What began largely as a voluntary practice is now being incorporated into formal climate-governance requirements. The European Union's Corporate Sustainability Reporting Directive requires companies covered by the Directive to obtain external assurance over their sustainability reporting [4, 5]. At the same time, the Science Based Targets initiative's Corporate Net-Zero Standard Version 2.0 introduces independent third-party assurance of target-base-year emissions inventories for Category A companies [6]. These requirements reflect a growing demand for credible emissions data to evaluate corporate climate commitments and transition plans.

The premise behind these requirements is that independent scrutiny makes corporate emissions inventories more reliable. Existing studies associate assurance with higher-quality environmental and carbon disclosure [2, 7, 8], systematic differences in reported

emissions [9], and a lower cost of equity [10]. Evidence on carbon performance is more mixed. Higher levels of assurance have been associated with improved reported carbon performance [11], while recent evidence finds that external assurance strengthens the relationship between credible climate transition plans and subsequent carbon performance [12]. Self-reported corporate emissions are also subject to substantial ex post revisions [13], making it difficult to distinguish changes in reported inventories from changes in firms' underlying emissions.

However, most of this evidence does not isolate the causal effect of verification, making it difficult to separate the effect of verification from differences between firms that obtain it and those that do not. Recent work strengthens identification by exploiting cross-country variation in mandatory assurance under the EU Non-Financial Reporting Directive and finds improvements in a model-based measure of Scope 1 and Scope 2 carbon-accounting quality [14]. What remains unclear is whether verification changes reported emissions themselves, whether these effects differ across scopes, and how they evolve after adoption.

Scope 3 presents a particularly difficult measurement problem because firms must determine which categories are relevant and estimate emissions across complex value chains [15, 16]. The closest evidence on behavioral effects related to Scope 3 comes from supply chains, where customer Scope 3 disclosure is associated with lower supplier emissions, with a stronger association when those disclosures are externally assured [17]. This evidence, however, concerns suppliers rather than the disclosing firm's own emissions inventory.

More importantly, existing evidence does not distinguish between two ways in which verification could affect corporate emissions. Verification may have a reporting effect by identifying omissions, inconsistent boundaries or other measurement errors, changing reported emissions around the verification process. But verification may also have a behavioral effect if the information generated through external scrutiny changes subsequent managerial decisions. More accurate inventories can reveal where emissions arise. This could help identify previously overlooked sources and inform decisions about efficiency, procurement or abatement. These responses should take time to affect emissions, unlike reporting adjustments associated with preparing and verifying the inventory. This distinction raises two empirical questions. How does verification change reported emissions across scopes around adoption? Does it also change subsequent firm behavior?

We exploit the staggered adoption of third-party verification among firms reporting to CDP between 2018 and 2025. Because firms generally verify their emissions inventory as a bundle, we track changes across Scope 1, Scope 2 and Scope 3 around the first adoption of verification. Using a staggered difference-in-differences design [18], we compare treated firms with never-verified firms, estimating both the adoption-year effect and its evolution over time. Comparing the timing of these effects across emissions scopes helps separate changes in reporting from changes in firm behavior. Effects that appear right at adoption are more consistent with a reporting effect, correcting information firms already had.

Effects that build up over several years are more consistent with a behavioral effect, reflecting real shifts in firm behavior.

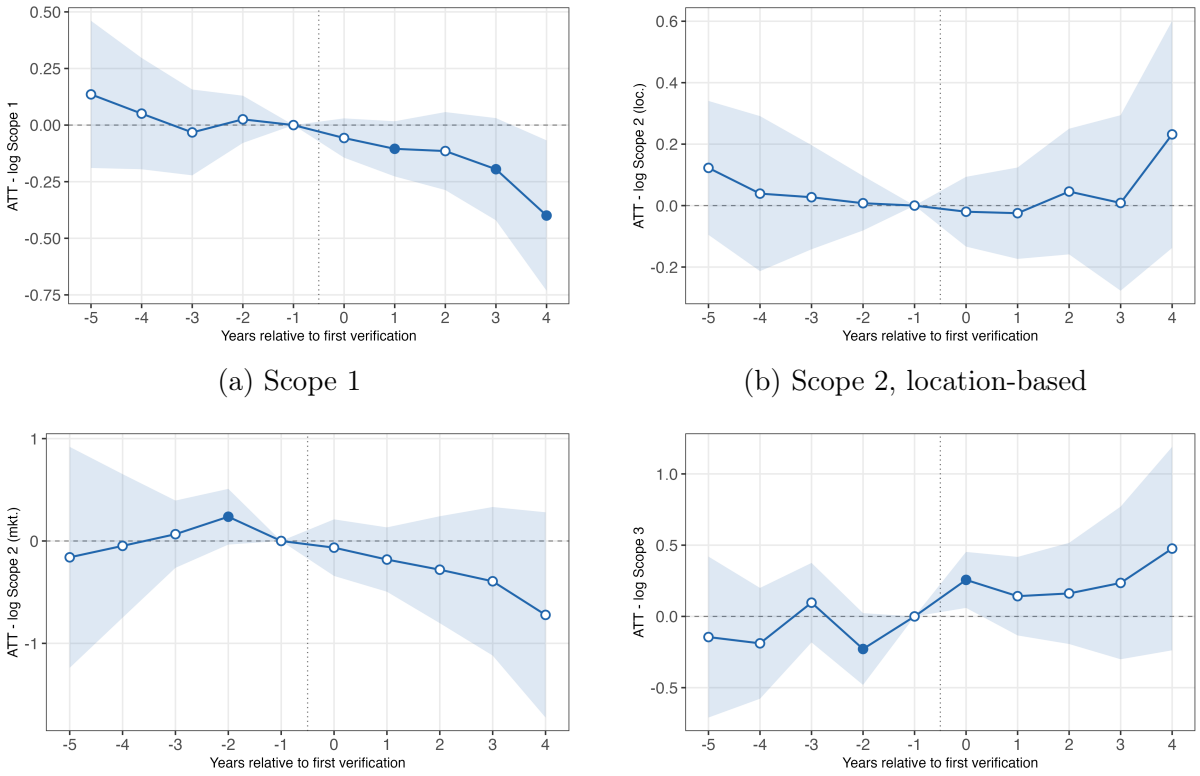
Results

Emissions responses differ across scopes

Figure 1 shows different effects across emissions scopes. Pre-treatment estimates are generally centered around zero, with no clear evidence of systematic differences before verification adoption. Figure 1d shows that reported scope 3 emissions increase sharply, by approximately 29 percent, in the first year of verification, after which the point estimates remain positive but show no comparable additional jump. The timing of this increase points to a reporting effect in which verification changes the coverage and measurement of value-chain emissions. Scope 1 follows a different pattern as seen in Figure 1a. Reported emissions do not change when verification is first introduced, but decline progressively in subsequent years. The reduction becomes larger further from adoption, averaging approximately 11 percent across the post-adoption period. There is no adjustment right when firms adopt verification. Emissions then fall by more each year, which looks more like a delayed behavioral response than a one-time correction. By contrast, neither location-based nor market-based Scope 2 emissions show a robust change following verification (Figures 1b and 1c). One explanation is that purchased electricity is comparatively easy to measure from meters, invoices and contractual information, leaving less scope for verification to uncover large errors. For location-based emissions, firms also have limited control over the carbon intensity of the electricity grid. Electricity consumption itself, which firms can control, also does not change significantly after verification, ruling out an efficiency-driven decline as well. The absence of a comparable Scope 2 response also argues against a general tendency for reported emissions to change mechanically after verification is adopted.

Fuel use falls while total energy does not

To explain Scope 1, Figure 2a shows that verification is followed by a gradual decline in firms' absolute fuel consumption, while total energy consumption does not change. The response is small around adoption but becomes larger in subsequent years, matching the delayed decline in Scope 1 emissions. Fuel use therefore falls without a broader decline in energy consumption. The share of fuel in total energy follows the same gradual pattern (Fig. 3). Because fuel combustion is a direct source of Scope 1 emissions, the joint decline in Scope 1 emissions and fuel use suggests that firms reduce their reliance on fuels after verification rather than simply revising their reported Scope 1 inventory. This could reflect



(a) Scope 1

(b) Scope 2, location-based

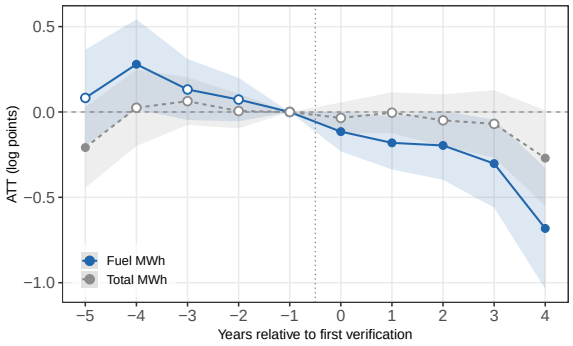
(c) Scope 2, market-based

(d) Scope 3

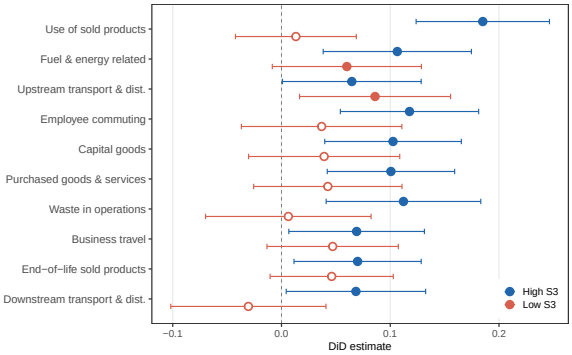
Figure 1: Effect of third-party verification on reported emissions. Each panel shows Sun-Abraham staggered difference-in-differences estimates of the effect of first-year verification adoption on log-scale reported emissions, relative to the year before first verification. The sample covers CDP-reporting firms from 2018 to 2025. Control firms are never-treated firms whose panel spans at least one year before and one year after the earliest treated cohort's adoption year. Shaded bands are 95% confidence intervals. Solid dots indicate estimates statistically significant at least at the 10% level and hollow dots are non-significant.

fuel substitution, efficiency improvements, or other operational changes. This matches qualitative evidence that assurance often involves recommendations to management and the identification of operational issues that change business practices [19]. Related evidence shows that third-party-certified environmental management systems can reduce operational carbon emissions [20]. We cannot determine whether the fuel response comes from verifier recommendations, better internal information or other changes triggered by assurance. The fact that fuel consumption itself falls, and does so gradually after adoption, points to a behavioral rather than purely reporting response. An unobserved factor could in principle drive both verification adoption and the later Scope 1 decline. Three patterns argue against this. Pre-treatment trends are flat, unlike the Scope 3 extensive margin, with no sign of early reductions. The decline starts only after adoption and grows over time, consistent with a slow operational response rather than an immediate confound. The effect is also specific to fuel use, with no matching change in Scope 2. This does not

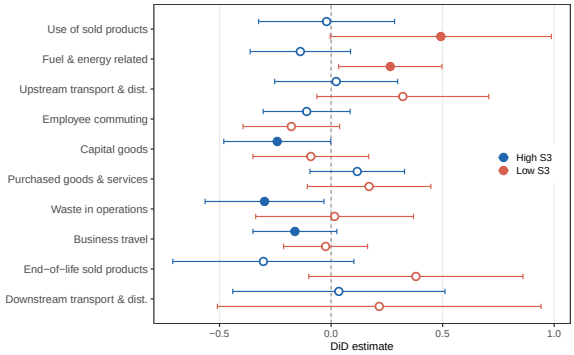
rule out an omitted factor, but favors a channel tied to verification.



(a) Fuel and total energy consumption



(b) Scope 3 Extensive margin



(c) Scope 3 Intensive margin

Figure 2: Mechanisms: fuel consumption and Scope 3 category reporting. Panel (a) shows Sun-Abraham staggered difference-in-differences estimates of the effect of third-party verification on log fuel consumption and log total energy consumption. The solid blue line reports fuel MWh and the dashed grey line total MWh. Estimates use the same never-treated spanning controls and 2+2 restriction as Figure 1. Shaded bands are 95% confidence intervals. Panels (b) and (c) show DiD estimates for the extensive margin (an indicator equal to one if the firm classifies the category as Relevant, calculated) and intensive margin (log emissions per reported category), respectively. Estimates are shown separately for firms with pre-treatment Scope 3 emissions above and below the sample median. In all panels, solid dots indicate estimates statistically significant at least at the 10% level and hollow dots are non-significant. Standard errors are clustered at the firm level throughout.

Scope 3 changes operate through category reporting

For Scope 3 emissions, the evidence points instead to changes in reporting. Category-level estimates show changes both in whether individual Scope 3 categories are reported and in the emissions recorded within already reported categories (Figures 2b and 2c). These responses vary across categories and between firms with higher and lower pre-verification Scope 3 emissions. On the extensive margin, whether a category is reported at all, higher-emitting firms expand coverage across several categories, while changes among

lower-emitting firms are smaller. Category coverage also begins expanding before formal adoption, suggesting that firms start building their inventories before verification. This could reflect an anticipation effect, where firms prepare their inventories for external scrutiny before formal adoption, or firms already improving their inventories being more likely to pursue verification. We therefore interpret these pre-adoption changes as part of an inventory-building process around verification rather than an effect of formal verification itself.

The intensive margin, how much is reported within categories already covered, shows a different pattern. Among lower-emitting firms, verification leads to substantial upward revisions in some already-reported categories, including an increase of approximately 65 percent in reported emissions from the use of sold products. Among higher-emitting firms, reported emissions instead fall in several categories, including capital goods, waste in operations and business travel. Verification therefore does not simply uncover previously unreported emissions. It also changes both what enters the inventory and the values assigned to emissions already being reported, with corrections occurring in both directions. The aggregate Scope 3 increase reflects the net effect of these corrections. Fig. 4 shows the extensive and intensive margins separately for high- and low-Scope 3 firms.

Discussion

An increase in reported emissions following verification need not indicate worsening climate performance. For Scope 3, it may instead reflect a more complete and potentially more accurate inventory, with some improvements in coverage beginning before formal verification. This matters for climate targets measured against a reported baseline, since an incomplete baseline distorts both initial footprints and subsequent reduction estimates. The delayed declines in Scope 1 emissions and fuel consumption, with no comparable decline in total energy consumption, point to changes beyond reporting. Our design does not identify the precise mechanism, but the pattern is consistent with prior evidence that assurance can generate recommendations and identify operational improvements that influence subsequent firm behavior. These results suggest that third-party verification plays two distinct roles in corporate climate governance. Where emissions are difficult to measure, verification can reshape the reported inventory. Where emissions are tied more directly to operations, it may also influence subsequent firm behavior. This distinction is important for interpreting verified emissions data, evaluating corporate climate targets, and designing emerging assurance requirements. For firms, verification may offer value beyond external credibility by improving the information available for internal emissions management and operational decision-making.

Methods

Data

Our data come from the Carbon Disclosure Project (CDP) annual GHG emissions dataset from 2018 to 2025. For each firm-year, CDP records self-reported scope 1 direct emissions, scope 2 location-based and market-based purchased energy emissions, and the total value-chain emissions across the 15 standard GHG Protocol Scope 3 categories. For scope 3 emissions, CDP also records whether each category is classified as "Relevant, calculated," "Relevant, not yet calculated," "Not relevant, explanation provided," "Not relevant, calculated," or "Not evaluated," along with the reported emission value in tCO_2e where applicable. CDP records the scope of any third-party verification process the firm has in place, which is our treatment variable. CDP records this field as "third-party verification or assurance." We use the term verification to refer to this treatment variable throughout our empirical analysis, and retain the terminology of specific standards or studies where we describe their own requirements or findings. For the fuel mechanism, CDP reports energy consumption in MWh separately for fuel combustion, purchased electricity, heat, steam, cooling, and self-generation. We define total energy consumption as the sum of these sources and estimate effects on fuel consumption and total energy consumption separately. We additionally examine fuel share, defined as fuel combustion divided by total energy consumption, as a supplementary outcome.

Sample construction

The 2019 public CDP Scope 3 workbook covers an unusually small number of respondents. We exclude 2019 from all models to maintain a consistent sample window across scopes and use years 2018 and 2020-2025. Because the data begin in 2018 and 2019 is excluded, treated firms must have first verified in 2021 or later to satisfy the requirement of at least two pre-treatment observations. We also exclude the 2020 first-adoption cohort because COVID-19 disrupted both energy use and value-chain activity in that year, confounding the verification effect across all scopes. Treated firms must have at least two pre-treatment and two post-treatment observations with non-missing outcome data for the relevant scope. Control firms must satisfy the spanning restriction (at least one observation $\leq$ 2020 and at least one $\geq$ 2021). The Scope 3 sample includes 460 treated firms and 407 spanning controls. Scope 1 and location-based Scope 2 samples include 452 and 422 treated firms respectively, with 481-564 controls. Market-based Scope 2 is smaller (265 treated, 233 controls) due to incomplete market-based reporting. The category-level Scope 3 mechanism uses the same restrictions, with treated firms split at the median of pre-treatment Scope 3 emissions (230 firms in each group, all 407 controls). As a robustness check, we re-estimate the Scope 1 and Scope 2 models on a broader panel

that includes all firms in the CDP verification records regardless of Scope 3 reporting completeness, expanding the Scope 1 treated sample to 668 firms. Results are nearly identical.

Empirical strategy

We estimate the causal effect of first-year verification adoption using the Sun-Abraham interaction-weighted estimator, which avoids the heterogeneous treatment-effect bias present in two-way fixed-effects regressions under staggered adoption. All specifications include firm, year, industry-by-year, and country-by-year fixed effects, and control for pre-treatment climate incentive status interacted with year. Standard errors are clustered at the firm level.

Outcome variables

The primary outcome is $\log(1 + \text{emissions})$ in tCO_2e. For the Scope 3 category analysis we construct an extensive margin indicator equal to one if a category is classified *Relevant, calculated*, and an intensive margin equal to log emissions per reported category. As an additional robustness check, we re-estimate all the emissions event studies on a propensity-score matched sample (nearest-neighbor matching on industry, region, pre-treatment outcome level, and pre-treatment climate incentive status). Results are consistent with the main specification.

References

[1] Roger Simnett, Ann Vanstraelen, and Wai Fong Chua. Assurance on sustainability reports: An international comparison. *The Accounting Review*, 84(3):937–967, 2009.

[2] Brandon Gipper, Samantha Ross, and Shawn X Shi. Esg assurance in the united states. *Review of Accounting Studies*, 30(2):1753–1803, 2025.

[3] Rina Datt, Pranil Prasad, Connie Vitale, and Krishan Prasad. International evidence of changing assurance practices for carbon emissions disclosures. *Meditari Accountancy Research*, 30(6):1594–1628, 2022.

[4] European Parliament and Council of the European Union. Directive (EU) 2022/2464 of the european parliament and of the council of 14 december 2022 as regards corporate sustainability reporting. Official Journal of the European Union, L 322, pp. 15–80, 2022. URL `https://eur-lex.europa.eu/eli/dir/2022/2464/oj`. Corporate Sustainability Reporting Directive.

[5] European Parliament and Council of the European Union. Directive (EU) 2026/470 of the european parliament and of the council of 24 february 2026 amending directives 2006/43/EC, 2013/34/EU, (EU) 2022/2464 and (EU) 2024/1760 as regards certain corporate sustainability reporting requirements and certain corporate sustainability due diligence requirements. Official Journal of the European Union, L 2026/470, 2026. URL https://eur-lex.europa.eu/eli/dir/2026/470/oj.

[6] Science Based Targets initiative. Corporate net-zero standard version 2.0. Technical report, Science Based Targets initiative, London, June 2026. URL https://files.sciencebasedtargets.org/production/files/Corporate-Net-Zero-Standard-version-2.pdf.

[7] Robyn Moroney, Carolyn Windsor, and Yong Ting Aw. Evidence of assurance enhancing the quality of voluntary environmental disclosures: An empirical analysis. *Accounting & Finance*, 52(3):903–939, 2012.

[8] Le Luo, Qingliang Tang, Hanlu Fan, and Jamie Ayers. Corporate carbon assurance and the quality of carbon disclosure. *Accounting & Finance*, 63(1):657–690, 2023.

[9] Georgios Papadopoulos. The effect of external assurance on corporate carbon disclosures: Empirical evidence from europe. *Energy Economics*, 153:109084, 2026. doi: 10.1016/j.eneco.2025.109084.

[10] Rina Datt, Le Luo, and Reuben Segara. Voluntary carbon assurance and the cost of equity capital: International evidence. *Australian Journal of Management*, 50(4): 1129–1165, 2025.

[11] Alireza Rohani, Mirna Jabbour, and Sulaiman Aliyu. Corporate incentives for obtaining higher level of carbon assurance: seeking legitimacy or improving performance? *Journal of Applied Accounting Research*, 24(4):701–725, 2023.

[12] Mega Silvia, Deyan Zhao, Donghuan Bai, Dong Xiang, Dewi Sartika, Andrian Noviardy, Fei Guo, Jing Jin, Tangqing Song, Tianyi Cao, et al. Do credible climate transition plans matter for carbon performance? evidence from fortune global 500 firms. *Frontiers in Environmental Science*, 14:1907357, 2026.

[13] Lauren Cohen, Ethan Rouen, and Kunal Sachdeva. Widespread revisions of self-reported emissions by major us corporations. *Nature Climate Change*, 16(1):33–36, 2026.

[14] Brandon Gipper, Fiona Sequeira, and Shawn X. Shi. Carbon accounting quality: Measurement and the role of assurance. *Journal of Accounting and Economics*, 81 (2):101849, 2026. doi: 10.1016/j.jacceco.2025.101849.

[15] Christian Blanco, Felipe Caro, and Charles J Corbett. The state of supply chain carbon footprinting: analysis of cdp disclosures by us firms. *Journal of Cleaner Production*, 135:1189–1197, 2016.

[16] Lena Klaaßen and Christian Stoll. Harmonizing corporate carbon footprints. *Nature Communications*, 12(1):6149, 2021. doi: 10.1038/s41467-021-26349-x.

[17] Young Jun Cho, Jungbae Kim, Holly Yang, and Mengjie Yang. Scope 3 emissions disclosure and supply chain decarbonization. Research Paper 2025-188, Singapore Management University, School of Accountancy, 2024.

[18] Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of econometrics*, 225(2):175–199, 2021.

[19] Charika Channuntapipat. Can sustainability report assurance be a collaborative process and practice beyond the ritual of verification? *Business Strategy and the Environment*, 30(2):775–786, 2021.

[20] Abdoul G. Sam and Danbee Song. Iso 14001 certification and industrial decarbonization: an empirical study. *Journal of Environmental Management*, 323:116169, 2022.

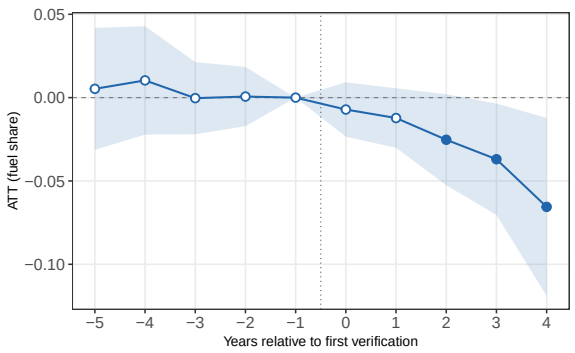
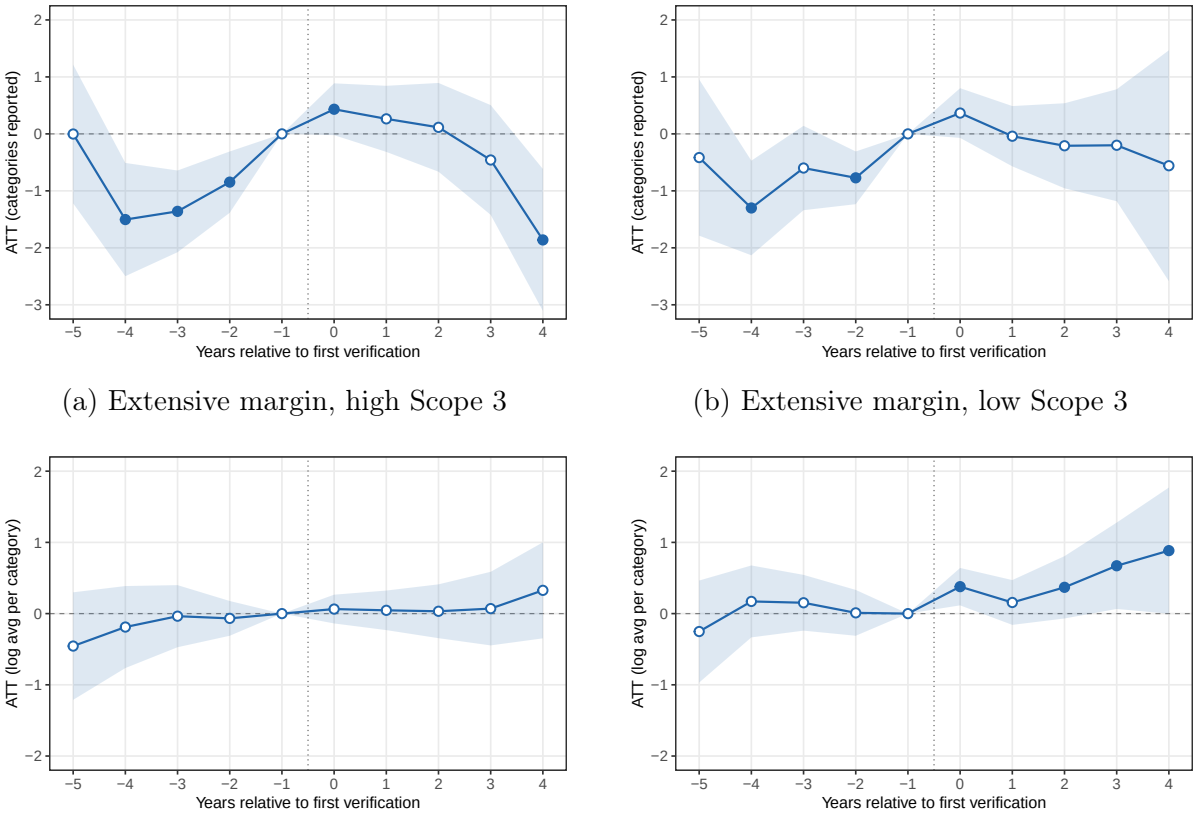


Figure 3: Fuel share following third-party verification. Sun–Abraham staggered difference-in-differences estimates of the effect of first verification on the share of total energy consumption from fuels. Shaded bands are 95% confidence intervals. Solid dots indicate estimates statistically significant at least at the 10% level and hollow dots are non-significant.



(a) Extensive margin, high Scope 3

(b) Extensive margin, low Scope 3

(c) Intensive margin, high Scope 3

(d) Intensive margin, low Scope 3

Figure 4: Dynamic effects of third-party verification on Scope 3 reporting margins. Panels (a) and (b) show Sun-Abraham staggered difference-in-differences estimates for the extensive margin. Panels (c) and (d) show estimates for the intensive margin. High and low Scope 3 groups are defined according to whether firms' pre-treatment Scope 3 emissions are above or below the sample median. Estimates are relative to the year before first verification. Shaded bands are 95% confidence intervals. Solid dots indicate estimates statistically significant at least at the 10% level and hollow dots are non-significant. Standard errors are clustered at the firm level.