Overestimated until disclosed? Vendor emission estimates and firms' first GHG emission disclosures

Draft version: November 2026

Abstract

Data vendors fill gaps in corporate emission disclosure with model-based estimates, and investors and researchers frequently pool these estimates with firm-reported figures. As there is evidence on systematic differences between estimated and recorded emissions, we quantify what happens to a firm's recorded emissions when the vendor LSEG stops estimating and using the firm's own published figure. Using LSEG data on 10,067 non-financial firms from 2015 to 2023 and a staggered difference-in-differences design, we find that recorded Scope 1 and Scope 2 emissions fall on average by 60% and 51% respectively in the year of the switch, relative to firms that are still estimated. The drop is consistent with a change in measurement rather than a real emission reduction. Firms do not switch mainly because they were overestimated.

Keywords: GHG emission disclosure, ESG data, vendor estimates, measurement error, staggered difference-in-differences

JEL classification: G14, M41, Q54, Q56

1. Introduction

Firm-level emissions data have become an input to capital allocation (Krueger et al., 2020). Asset managers screen portfolios on carbon intensity, banks report financed emissions, and researchers price carbon risk (Bolton et al., 2021) and evaluate corporate emission reduction policies with such data (König-Sykorova, 2026). Emission data is also needed to track whether finance flows are consistent with the Paris Agreement (Zamarioli et al., 2021). As many firms, however, do not track and disclose their emissions (Aswani et al., 2024), data vendors fill gaps with model-based estimates (LSEG, 2025). For instance, in Compustat, only 5.3% of firm-years carry an emissions figure that the firm itself disclosed (El Ghoul et al., 2026).

While there is a growing interest in GHG emission data verification (Source), users of emission data tend to see a number per firm-year, with a method flag that is easy to ignore. Also many academic studies use estimated and reported values together (Aswani et al., 2024; El Ghoul et al., 2026). Recorded GHG emissions tend to fall when a firm's data change from estimated to disclosed (El Ghoul et al., 2026), whereas studies that correct for selection into disclosure conclude that non-reporting firms emit more than estimates based on reporters suggest (Crippa et al., 2025; Chen et al., 2025). The fall at first disclosure may reflect a general upward bias in vendor estimates, or selection.

We use the first year in which LSEG records a firm's own Scope 1 and Scope 2 figures, respectively, and compare it with the year before, where data is based on model estimates. The change in the recorded value at the switch shows how far the estimate was from the firm's own figure, which we treat not as the true value, but as the firm's own measure. We compare 1,367 firms that switch, with firms that are still estimated in the same years, using the difference-in-differences estimator of Callaway and Sant'Anna (2021).

Three results stand out. First, recorded Scope 1 and Scope 2 emissions fall by 60% and 51% respectively in the year of the switch. The drop narrows to 48% and 46% lower three years later. Second, this drop is consistent with a change in measurement rather than a real reduction. Concretely, a one-year drop of 50% or more occurs in 52% of switches from estimated to reported, but in only 3% to 5% of firm-years in which the data source remains estimated or reported respectively. In addition, firms' total assets, revenue and investment do not shrink at the switch. They grow slightly faster than in firm-years in which the data source does not change. Third, firms do not appear to switch mainly because they were overestimated. Firms that start to report are only slightly cleaner than the firms that already report in their industry. Even a firm whose reported emission intensity is at the industry average had been estimated at

about twice the value it then reported, and one in three switchers reports even more than the vendor had estimated.

We contribute to research on the quality of emissions data by measuring, within firms, the direction and size of the difference between vendor estimates and firms' own figures, and by showing that this difference is not confined to the firms that had a reason to report. For users, a change in the method flag from estimated to reported is a change of measure, not a sign of decarbonisation. Our evidence covers one vendor and firms that start to report at some point; whether estimates are also too high for firms that never report cannot be observed directly.

2. Data

We use LSEG's Scope 1 and Scope 2 emissions in tonnes of CO_2 equivalent, which contain a flag for its origin. LSEG collects reported values from annual and sustainability reports and CDP responses and screens them against industry bounds. A reported value outside these bounds is capped to the nearest bound.[1] Where both exist, LSEG uses the value from the firm's own reports rather than from CDP. For Scope 2, it uses location-based values where available. If a firm reports the sum of Scope 1 and 2 and one of the two, LSEG derives the other as the difference. A reported Scope 1 or 2 value is not replaced by the model estimate just because it is lower (LSEG, 2025). When no usable report exists, LSEG estimates emissions as the median of three models. First, a sector-median model, second, a model based on the firm's business segments, and third, an environmentally extended input-output model (LSEG, 2025). Each model multiplies an emission intensity by revenue, so the estimate scales with sales. The models are informed by reporting firms in the FTSE All-Cap index.

The original data covers firm-year observations between 2015 and 2025. We drop 2024 and 2025, for which coverage is still incomplete. We exclude firms in Finance and Insurance. The final sample contains 56,130 firm-years of 10,067 firms.

We define the switch as the first year in which a firm's own figure enters the series, including values that LSEG caps[2], provided that the firm carried a model estimate in the year before. The switch is thus the first year in which LSEG records a value reported by the firm. The first

[1] LSEG calls this winsorizing. It compares the firm's reported emissions per unit of revenue with those of its industry peers: reporting firms that LSEG uses to train its models, in the same ICB subsector or, if it has fewer than 20 such firms, the next broader ICB level, over a rolling three-year window. For Scope 1 and 2, a log intensity more than 1.0 interquartile range below the first quartile or above the third quartile of the peers counts as extreme; for Scope 2, LSEG sets the bounds with location-based values only. LSEG first carries forward an earlier, non-extreme intensity of the firm, by up to two years, and sets the remaining extreme values to the bound, flagged as "Winsorized_low" or "Winsorized_high". A first disclosure has no earlier value of the firm, so an extreme first value is capped. Extreme values taken from CDP are not capped but dropped (LSEG, 2025).

[2] In total LSEG caps 10% of first values because they are extreme for the industry, three quarters of them because they are very low, and excluding them would drop the firms whose first values differ most from the estimate

reported value may coincide with the firm's first year of publication or appear later, as data collection, ingestion and verification by the vendor lag behind. The data do not show whether the switch year marks the firm's first publication or only the vendor's first record of it. Because we use each firm's first disclosure, every switch belongs to a different firm. This yields 2,870 switches between 2016 and 2023. Manufacturing accounts for 38% of them, and firms headquartered in the United States and China for 42% (Annex Tables A2 and A3).

Table 1 shows how we arrive at the event-study sample. Firms that never disclose (1,471) remain in the sample as the never-treated comparison group. Firms without a clean switch, that is, firms that report from their first observation or revert to estimates after disclosing, drop out.. The event-study sample contains 20,263 firm-years of 2,838 firms, incl. 1,367 switchers and 1,471 firms that never disclose.

Table 1: Sample selection

	Firm-years	Firms
All firm-years with an LSEG Scope 1 value, 2015–2023, excl. Finance and Insurance	56,130	10,067
Firms that already report in their first year in the data	−22,158	−3,949
Firms whose first disclosure follows an extrapolated or energy-model value or a gap in the data	−792	−113
Switchers of 2021–2023 (fewer than 3 yrs observed after switch)	−7,262	−1,314
Switchers of 2016–2020 whose value is later again estimated or extrapolated	−1,490	−189
Firm-years not observed in both years of a comparison	−4,165	−1,664
Event-study sample	**20,263**	**2,838**
of which switchers of 2016–2020		1,367
of which firms that never disclose		1,471

Each row removes firm-years that cannot be used in the event study (Table 3, column 1). Firms that never disclose stay in the sample throughout and serve as comparison firms. The first two rows remove firms without a switch from a model estimate to their own figure. An extrapolated value carries forward the firm's own earlier intensity for up to two years, either after the firm stops reporting or in place of a later value that LSEG considers extreme; the method flag does not show which (LSEG, 2025). The last row removes firm-years that the estimator cannot use because the firm is not observed in both years of a comparison. The main specification with control variables (Table 3, column 3) uses 17,957 of these firm-years, because some firms lack financial data.

Table 2: Descriptive statistitics

	Never disclose	Switchers, year before switch	Difference (log or level)	t
Scope 1 emissions, estimated (t)	3,038	27,147	2.19	38.1***
Scope 1 intensity, estimated (t per USD m revenue)	14.5	25.8	0.58	16.5***
Scope 2 emissions, estimated (t)	5,974	38,473	1.86	38.1***
Scope 2 intensity, estimated (t per USD m revenue)	27.8	35.7	0.25	13.1***
Revenue (USD m)	207	1,064	1.64	35.6***
Total assets (USD m)	558	2,102	1.33	38.7***
Return on assets	−0.068	0.040	0.108	13.9***
PP&E / total assets	0.218	0.278	0.060	12.7***
Total debt / total assets	0.252	0.263	0.012	2.0**

Both groups carry LSEG model estimates in the year shown. Scope 1 rows compare the groups defined by the Scope 1 series, Scope 2 rows the groups defined by the Scope 2 series. Emissions, intensity, revenue and total assets are geometric means (exponential of the mean of logs); the difference and t-statistic refer to the logs. Other rows are arithmetic means. t-statistics from two-sample t-tests; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Other variables have slightly fewer observations.

In the year before their switch, switchers are larger, more profitable and more capital-intensive than firms that never disclose, in line with evidence that larger firms are more likely to report emissions (Crippa et al., 2025). Their estimated Scope 1 and Scope 2 emissions are also higher (Table 2). Because the estimate scales with revenue, this partly reflects their size.

3. Theoretical and empirical literature

Disclosure theory predicts that firms with good news disclose and that silence is read as bad news, unless disclosure is costly or investors are unsure whether the firm has the information (Grossman, 1981; Milgrom, 1981; Verrecchia, 1983; Dye, 1985). Standardised frameworks such as the GHG Protocol and CDP have lowered the cost of reporting, and investors increasingly ask for climate information (Flammer et al., 2021; Ilhan et al., 2023). Disclosure nevertheless remains selective. High emitters may stay silent to pool with other non-disclosers (Matsumura et al., 2014), and voluntary registries allow firms to report favourable figures (Kim and Lyon, 2011). Visibility works in the other direction. Firms that are larger or cause more environmental damage attract more scrutiny from regulators and the public, and they disclose more substantively (Marquis et al., 2016).

Vendor estimates change this setting. Disclosure theory shows that information from outside the firm, such as market news or analyst reports, changes whether and when firms disclose; bad outside news, for example, can prompt a firm with better private information to disclose at once (Acharya et al., 2011; Frenkel et al., 2020). A vendor estimate is an outside information, but users may see it in place of the firm's own figure. A silent firm no longer pools with other non-disclosers but receives a specific number from the vendor's model. A firm that knows its emissions can compare them with this number. Because the estimate does not use the firm's private information, firms can select on it. This presumes, however, that firms know their emissions before they publish them. This is less plausible for firms that have never measured them, in particular smaller firms, for which the fixed costs of measuring and reporting emissions may be too high (Bolton et al., 2026). Chen et al. (2025) suggest that firms may disclose to correct a vendor's overestimate, but to our knowledge no study has tested this.

A second literature shows that estimated and reported emissions differ in ways that matter for inference. Reported Scope 1 and 2 figures align closely across providers, whereas estimates diverge, because each provider uses its own models, inputs and peer groups (Busch et al., 2022; El Ghoul et al., 2026). Estimates mainly reflect industry and size, so they miss differences between firms within an industry and identify the largest emitters less reliably (Kalesnik et al.,

2022). They are also highly correlated with financial fundamentals, which drives much of the association between emissions and stock returns (Aswani et al., 2024). These studies compare estimates across firms and treat the estimation error as noise; they do not show whether the estimates are, on average, too high or too low.

Whether estimates are too high or too low for non-reporters is less clear. Two recent studies model the decision to disclose and correct for it. Crippa et al. (2025) assume that the probability of disclosure depends on the firm's own, unobserved emissions, and Chen et al. (2025) use a Heckman-type selection correction combined with machine learning. Both conclude that non-reporters are likely to emit more than imputations based on reporters suggest, but their results depend on these assumptions about why firms disclose. El Ghoul et al. (2026) show for Trucost data that recorded Scope 1 emissions usually fall when a firm's data change from estimated to disclosed. They describe these one-year changes but do not compare them with firms that remain estimated, and they do not test whether the fall reflects real cuts or which firms choose to disclose. Reported values are not free of error either. Large US firms frequently revise them, and understatements are about twice as frequent as overstatements (Cohen et al., 2026). Our design therefore compares the vendor's measure with the firm's own measure. Thus, earlier work tells us that estimated and reported figures differ, but not how much the recorded value changes when a firm starts to report and whether this change reflects measurement rather than a real reduction. Nor does it tell us whether the change arises because only overestimated firms report, and this paper addresses both questions..

4. Hypotheses development

At the switch, both the source of the recorded value and, possibly, the firm's emissions may change. Firms might start to report after cutting emissions, for example to signal progress. Large real cuts, however, take time and usually require producing less, closing or selling plants, or investing in new equipment, which shows up in revenue, assets or investment. Running existing plants differently rarely halves direct emissions within a year, and even sustained carbon pricing is unlikely to deliver deep cuts quickly in hard-to-abate industries (Haites et al., 2024). A change of source, in contrast, moves the recorded value at once and only in the switch year.

H1: Recorded emissions change at the switch mainly because the source changes, not because emissions change.

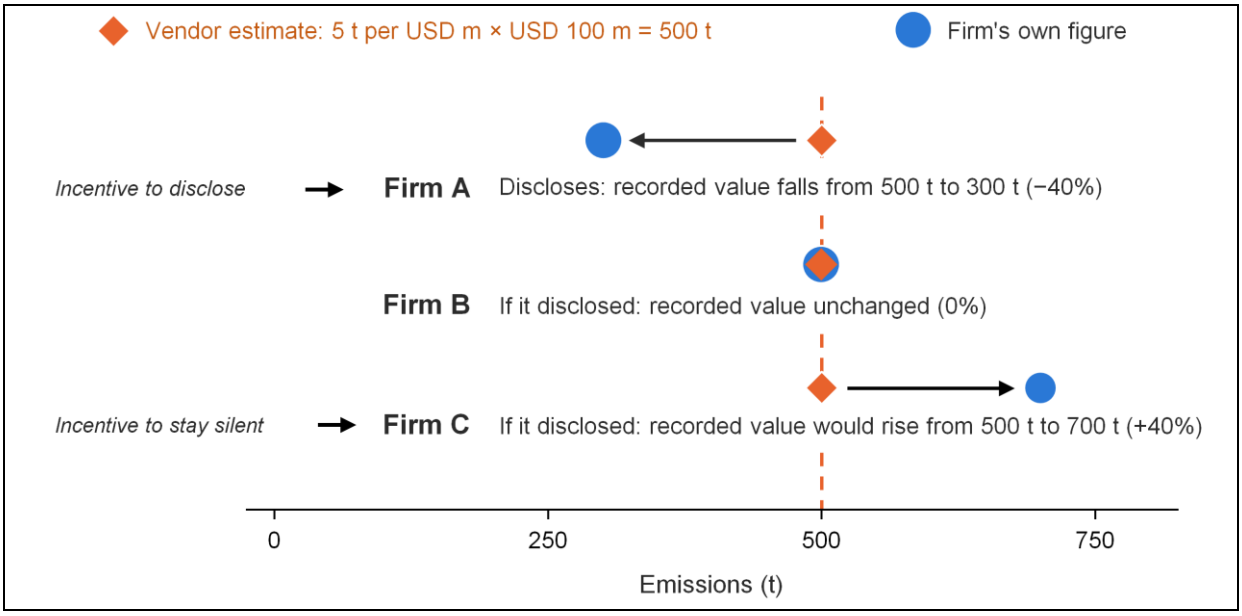
If H1 holds, the drop should occur exactly in the switch year and be much larger than the year-to-year changes of firms whose data source does not change, and revenue, assets and investment should not shrink at the switch.

H1 does not say why the first reported value lies below the estimate. Disclosure theory offers one answer (Verrecchia, 1983; Dye, 1985), where a firm that knows its emissions can compare them with the estimate it has been assigned. If its emissions are lower, disclosure lowers its recorded value; if they are higher, silence protects it. Firms would then select into disclosure when the estimate overstates them, and the drop would reflect who discloses rather than how the vendor estimates. Scrutiny, in contrast, would lead visible firms to disclose regardless of the estimate (Marquis et al., 2016). Figure 1 illustrates this logic for H2.

H2: Firms are more likely to disclose if the vendor's estimate overstates their emissions.

If H2 holds, first-time reporters should come from the clean end of their industry, and a firm whose reported emissions are typical for its industry should show little or no drop, because the estimate would be too high only for firms below the typical level.

Figure 1: Selective disclosure under a revenue-based vendor estimate



Hypothetical example. Firm A, B and C have the same revenue, USD 100 m, and do not report. Their true Scope 1 emissions are 300, 500 and 700 tonnes. As in a simplified version of the vendor's approach (LSEG, 2025), the estimate is an emission intensity of reporting peers multiplied by the firm's revenue: with a peer intensity of 5 tonnes per USD m, each firm is estimated at 500 tonnes (dashed line). Only A gains from disclosure, so its recorded value falls from 500 to 300. A drop is observed, but it reflects which firm discloses, and the switcher comes from the clean end of the industry.

5. Method

We estimate the drop with a difference-in-differences design. For a firm that switches in year g, we take the change in its recorded emissions from the year before the switch (g − 1), when the value is an estimate, to a later year t, and subtract the average change among firms that are still estimated in year t, that is, firms that switch later or never:

$$\Delta Y_i = \lambda_{gt} + \delta_{gt} \cdot D_i + \gamma_{gt} \cdot X_{i,g-1} + \varepsilon_i \quad (1)$$

$\Delta Y_i = Y_{i,t} - Y_{i,g-1}$ is the change in recorded absolute emissions, ln(GHG), D_i equals 1 for switchers, and γ contains the coefficients of the control variables. Callaway and Sant'Anna (2021) average these coefficients, weighted by the number of switchers. A single regression with firm and year fixed effects would also compare late switchers with firms that switched earlier, whose values have already dropped, and would understate the drop (Goodman-Bacon, 2021). The main sample includes firms that switch between 2016 and 2020, observed for three years after the switch. Firms whose value is later again estimated or extrapolated are excluded and added back in a robustness check (Table 1). The control variables $X_{i,g-1}$ include size (ln total assets), profitability (return on assets), capital intensity (PP&E / total assets) and sector dummies. Because equation (1) uses changes within firms and compares switchers with comparison firms in the same years, firm and year fixed effects drop out.

The estimator predicts the change each switcher would have had from comparison firms with similar values (Annex A.1). We do not use revenue, because LSEG computes its GHG emission intensity estimates from revenue figures. The key assumption is that, without the switch, the recorded emissions of switchers would have changed like those of firms that remain estimated. Because both groups carry vendor estimates before the switch, we can check this. In the years before the switch, the estimates of future switchers should not move differently.

We compare one-year changes in recorded Scope 1 and Scope 2 emissions at the switch with those in firm-years in which the source stays the same. In case the difference reflects real change, we expect to see this in changes in firm fundamentals, incl. total assets, revenue, property, plant and equipment (PP&E) and capital expenditure.

H2 is about which firms switch, and how far each switcher's first reported value lies from the estimate it would otherwise have received, which is called the gap. The gap of switcher i is its

change in ln(GHG) from $g-1$, the last estimated year, to g, the first disclosed year, minus the average change ΔY_k of the n_{jg} firms k in the same industry j that remain estimated in year g.[3]

$$Gap_i = \Delta Y_i - (1/n_{jg}) \Sigma_k \Delta Y_k \quad (2)$$

Subtracting the change of still-estimated firms removes movements that would have occurred anyway, for example because the vendor updates its model each year. A gap of zero means that the firm reports what the vendor would have estimated. A negative gap means that the estimate was too high, a positive gap means that the estimate was too low.[4]

To see which firms switch, we compute each switcher's entry rank, which uses only reported values. For each switcher, we take all firms that report emissions in its industry j in its switch year g, sort them by emission intensity and record the switcher's place r_i, from 1 (cleanest) to N_{jg} (dirtiest). These N_{jg} firms include firms that report for the first time in year g, like the switcher itself, and firms that have reported before; firms with estimated values are not included. Each switcher is thus compared only with firms of its own industry in its own switch year. The entry rank rescales the place to lie between 0 and 1. We compute it only where at least ten firms report in the industry and year:

$$Rank_i = (r_i - 0.5) / N_{jg} \quad (3)$$

The entry rank lies close to 0 for the cleanest and close to 1 for the dirtiest firm. Without selection, switchers would be spread evenly over their industry, and their mean entry rank would be 0.5.[5] Under H2, switchers come mainly from the clean end, so their mean entry rank lies clearly below 0.5. We test whether the mean entry rank differs from 0.5 (standard errors clustered by industry).

We then compare the gap of switchers with different entry ranks. If the vendor is right on average, switchers in the middle of their industry have a gap close to zero, and only switchers

[3] We do not observe the estimate a switcher would have received in its first disclosed year. LSEG (2025) describes how the estimate is built (Section 2), but not how the estimate of a given firm would have evolved. We therefore assume that it would have changed like the estimates of firms in the same industry that remain estimated, which share the same model updates and peer data. Because the models scale an emission intensity by revenue, a switcher whose revenue grows faster than that of these firms would have received a somewhat higher estimate. In the switch year, the revenue of switchers grows slightly faster than that of firms that remain estimated (10% against 9%, Table 5); allowing for this would make the gap about 0.01 log points more negative, so our measure is, if anything, conservative.

[4] For example, a firm estimated at 100,000 tonnes that reports 40,000 tonnes while estimates in its industry rise by 5% has a gap of $\ln(40{,}000/105{,}000) = -0.97$. Its first reported value is 62% below the estimate it would otherwise have received.

[5] Subtracting 0.5 places each firm in the middle of its place, so that the mean rank is exactly 0.5 without selection.

at the clean end report much less than estimated (see H2). If instead the estimate is too high for most firms that do not report, even switchers in the middle report less than the vendor had estimated.

6. Results

6.1. H1: Measurement error or real change

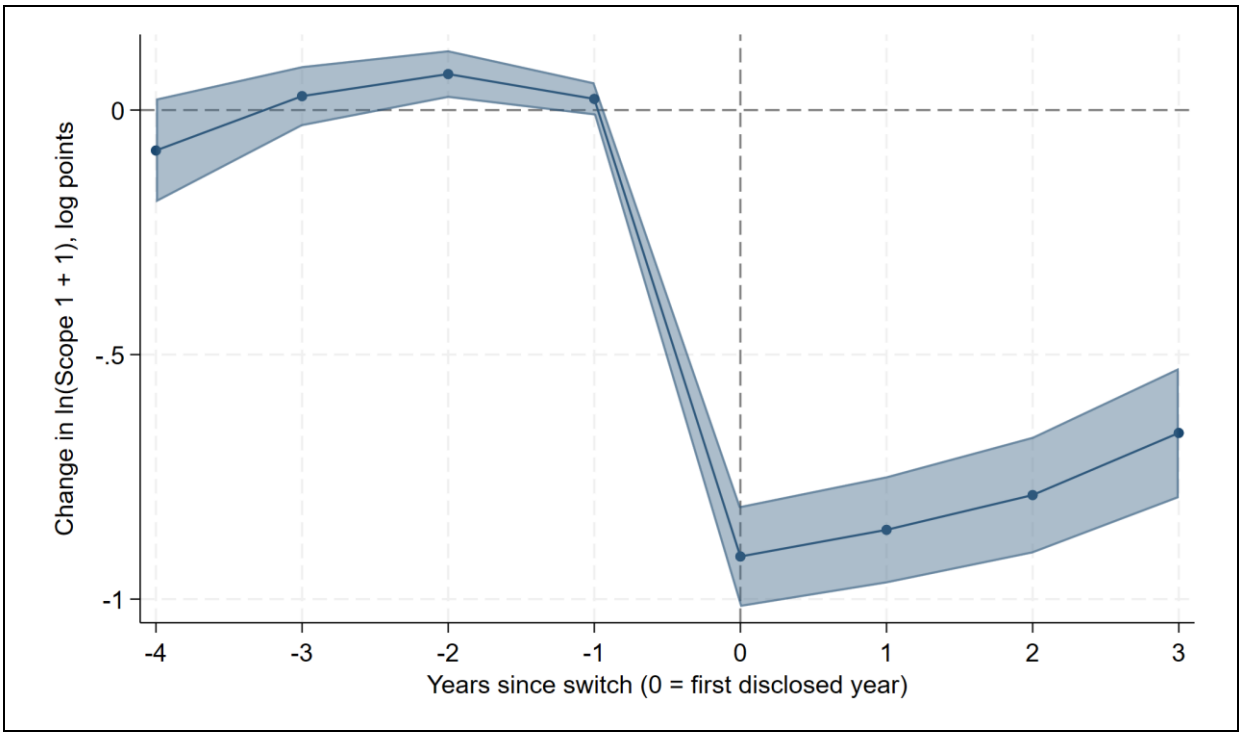
Table 3 and Figure 2 report the event-study estimates. Our main specification, with control variables, is shown in columns 3 and 4. Before the switch, recorded emissions of future switchers move largely, but not fully, in parallel with those of comparison firms. The pre-switch coefficients are small on average for Scope 1 and Scope 2 respectively, but a joint test rejects that they are all zero. The rejection comes from two single coefficients. Two years before the switch for Scope 1 and one year before for Scope 2, the estimates of future switchers rise by about 8% and 6% more than those of comparison firms. Without control variables, the joint test does not reject at the 5% level ($p = 0.087$ (0.44); columns 1 and 2). These deviations are small compared with the drop, and they point upward. Had the estimates of future switchers kept rising faster than those of comparison firms, the drop would be larger, not smaller. In the switch year, recorded Scope 1 (Scope 2) emissions fall by 60% (51%), relative to firms that remain estimated. The drop narrows in the following years but remains large. Three years later, recorded emissions are still 48% (46%) lower; in log points, the drop has shrunk by about 28% (14%), from 0.91 to 0.66 (0.72 to 0.62). The empirical design does not show why. Firms may revise low first values upwards or widen their reporting boundaries, or the estimates of comparison firms may change; we leave this question for future work. Measured as emission intensity (columns 7 and 8), the drop is similar, and the joint pre-trend test does not reject.

Table 3: Recorded emissions around the switch to self-reporting

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DV:	**Scope 1 (absolute)**	**Scope 2 (absolute)**	**Scope 1 (absolute)**	**Scope 2 (absolute)**	**Scope 1 intensity**	**Scope 2 intensity**	**Scope 1 intensity**	**Scope 2 intensity**
t − 4	−0.033	−0.015	−0.083	−0.012	−0.026	−0.012	−0.052	−0.012
	(0.046)	(0.041)	(0.054)	(0.045)	(0.027)	(0.020)	(0.032)	(0.020)
t − 3	0.030	0.023	0.028	0.022	0.018	0.004	0.022	0.013
	(0.027)	(0.026)	(0.031)	(0.030)	(0.016)	(0.015)	(0.019)	(0.016)
t − 2	0.057***	−0.004	0.074***	0.026	0.054***	−0.001	0.042**	0.004
	(0.022)	(0.019)	(0.025)	(0.020)	(0.016)	(0.012)	(0.019)	(0.014)
t − 1	0.006	0.022	0.023	0.054***	0.015	0.025**	−0.004	0.024**
	(0.015)	(0.014)	(0.017)	(0.016)	(0.011)	(0.010)	(0.013)	(0.012)
t (switch)	−0.968***	−0.708***	−0.913***	−0.719***	−0.974***	−0.714***	−0.950***	−0.755***
	(0.049)	(0.044)	(0.053)	(0.047)	(0.049)	(0.043)	(0.052)	(0.047)
t + 1	−0.974***	−0.729***	−0.858***	−0.685***	−0.953***	−0.700***	−0.924***	−0.742***
	(0.052)	(0.046)	(0.056)	(0.049)	(0.049)	(0.043)	(0.053)	(0.046)
t + 2	−0.915***	−0.719***	−0.787***	−0.664***	−0.911***	−0.714***	−0.916***	−0.767***
	(0.054)	(0.048)	(0.061)	(0.053)	(0.049)	(0.044)	(0.054)	(0.047)
t + 3	−0.803***	−0.671***	−0.660***	−0.617***	−0.811***	−0.677***	−0.816***	−0.736***
	(0.058)	(0.052)	(0.068)	(0.058)	(0.050)	(0.044)	(0.057)	(0.049)
Average pre-switch	0.015	0.006	0.011	0.022	0.015*	0.004	0.002	0.007
	(0.017)	(0.016)	(0.018)	(0.016)	(0.008)	(0.007)	(0.010)	(0.008)
Average post-switch	−0.915***	−0.707***	−0.805***	−0.671***	−0.912***	−0.701***	−0.902***	−0.750***
	(0.050)	(0.044)	(0.055)	(0.048)	(0.047)	(0.041)	(0.051)	(0.044)
Control variables	No	No	Yes	Yes	No	No	Yes	Yes
Pre-trend test: p-value	0.087	0.438	0.014	0.013	0.004	0.149	0.137	0.248
Observations	20,263	19,548	17,957	17,377	20,168	19,466	17,957	17,377

Callaway and Sant'Anna (2021) event-study estimates in log points. Dependent variable: ln(Scope 1 + 1) and ln(Scope 2 + 1) in columns 1–4, ln(Scope 1 / revenue) and ln(Scope 2 / revenue) in columns 5–8. Sample: firms switching 2016–2020 and firms that do not yet or never report, 2015–2023 (Table 1). Pre-switch coefficients compare each year with the previous year, post-switch coefficients with t − 1. Average pre-switch and average post-switch = mean of the coefficients from t − 4 to t − 1 and from t to t + 3. Pre-trend test = p-value of a joint test that all pre-switch coefficients are zero. Columns 3, 4, 7 and 8 (main specification) include the control variables of equation (1), measured in the year before the switch: ln(total assets), return on assets, PP&E / total assets and sector dummies; they use fewer firm-years because some firms lack financial data. Standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Further details see descriptive statistics (Table 2)

Figure 2: Recorded Scope 1 emissions around the switch to self-reporting



Callaway and Sant'Anna (2021) event-study estimates with 95% confidence intervals; see Table 3, column 3 (main specification). Pre-switch estimates compare each year with the previous year, post-switch estimates with the year before the switch (t − 1). Annex Figure A3 shows all estimates relative to t − 1.

Table 4 shows that drops of this size are rare when the data source does not change. Among firm-years in which a firm reports in both years, the median change in recorded Scope 1 emissions is zero. A decline of 50% or more occurs in 3.4% of these firm-years and in 5.3% of firm-years in which the firm is estimated in both years, but in 52% of switches. Real reductions of this size would therefore have to happen about ten to fifteen times more often exactly in the year in which a firm starts to report. The mean change at the switch, a fall of 63%, differs significantly from the rises of 3% and 4% in the other two groups. The switch also produces increases of 50% or more in 23% of cases, so the estimate is not too high for every firm, but it is too high far more often than too low. Scope 2 emissions show the same pattern (Panel B).

In addition, Table 5 shows that firms' operations do not shrink at the switch using total assets, revenue, PP&E and capital expenditure. In the switch year, total assets fall by 50% or more for only 0.2% of switchers, the same share as in firm-years in which the firm reports in both years (0.2%) and less than in firm-years in which it is estimated in both years (1.1%). Median growth of total assets at the switch (6%) is similar to that in other years (4% and 5%), while mean growth is a few percentage points higher (11%, against 6% and 9%).

Table 4: One-year changes in recorded emissions

Firm-years	N	Mean	Median	Drop ≥ 50%	Rise ≥ 50%	t vs. switchers
Panel A: Scope 1 (absolute)						
Reported in t − 1 and t	24,883	+3	0	3.4	9.4	28.3***
Estimated in t − 1 and t	16,442	+4	+4	5.3	12.8	28.5***
Switch: estimated in t − 1, disclosed in t	2,870	−63	−55	52.0	23.3	
Panel B: Scope 2 (absolute)						
Reported in t − 1 and t	24,077	+1	−1	3.0	7.7	24.7***
Estimated in t − 1 and t	16,493	+6	+4	4.7	13.2	26.0***
Switch: estimated in t − 1, disclosed in t	2,937	−53	−44	46.2	22.3	

Change in ln(emissions) from t − 1 to t, 2016–2023. Mean and median of the change, expressed as percentage changes (mean = 100 × (exp(mean change in logs) − 1)). t vs. switchers = t-statistic of a two-sample t-test (unequal variances) of whether the mean change in logs in the group differs from that at the switch; a positive value means that the values of the group rise more, or fall less, than at the switch; * p < 0.10, ** p < 0.05, *** p < 0.01. Drop ≥ 50% and rise ≥ 50% = share of firm-years (%) in which recorded emissions fall to half of the previous year's value or less, or rise by half or more. Groups are defined by the source of the Scope 1 value (Panel A) or the Scope 2 value (Panel B): reported = reported or capped value; switch = model estimate in t − 1 and first disclosure in t.

Revenue, PP&E and capital expenditure show the same pattern. In the year they start to report, switchers thus grow somewhat faster than in other firm-years, and most of these differences are significant. What matters for H1 is that their operations do not shrink. A real cut of Scope 1 emissions by more than half would in most cases require producing less, closing or selling plants, or large investments in new equipment, none of which is visible in the data. Cheaper changes, such as a switch of fuel, cannot be ruled out, but they would rarely halve direct emissions in the year a firm starts to report. Because the estimate scales a peer intensity by revenue, it also does not follow a firm's own cuts: cuts made in earlier years would appear only at the switch, as part of the distance between the estimate and the firm's own figure. Together, the evidence supports H1 as the drop is consistent with a change in measurement rather than a real reduction in emissions at the switch.

Table 5: One-year changes in firm fundamentals

Firm-years	N	Mean	Median	Drop ≥ 50%	Rise ≥ 50%	t vs. switchers
Panel A. Total assets						
Reported in t − 1 and t	24,851	+6	+4	0.2	3.3	−10.5***
Estimated in t − 1 and t	16,330	+9	+5	1.1	8.4	−3.8***
Switch: estimated in t − 1, disclosed in t	2,842	+11	+6	0.2	7.0	
Panel B. Revenue						
Reported in t − 1 and t	24,863	+6	+5	1.1	5.5	−6.8***
Estimated in t − 1 and t	16,342	+9	+7	3.4	10.8	−1.6*
Switch: estimated in t − 1, disclosed in t	2,842	+10	+8	1.2	8.2	
Panel C. Property, plant and equipment (PP&E)						
Reported in t − 1 and t	24,616	+7	+3	0.7	6.4	−7.0***
Estimated in t − 1 and t	15,790	+14	+5	1.8	14.5	0.4

Firm-years	N	Mean	Median	Drop ≥ 50%	Rise ≥ 50%	t vs. switchers
Switch: estimated in t − 1, disclosed in t	2,767	+13	+6	0.8	12.0	
Panel D. Capital expenditure						
Reported in t − 1 and t	24,784	+4	+4	5.9	16.6	−2.9***
Estimated in t − 1 and t	16,040	+5	+5	12.7	25.5	−2.3**
Switch: estimated in t − 1, disclosed in t	2,825	+8	+9	8.2	23.5	

Change in ln(total assets), ln(revenue), ln(PP&E) and ln(capital expenditure) from t − 1 to t, 2016–2023. Columns as in Table 4. Groups are defined by the source of the Scope 1 value, as in Table 4, Panel A. Capital expenditure in absolute value.

6.2. H2: Self-selection

Table 6 and Figure 3 report the results for H2. Under H2, firms disclose mainly if the vendor's estimate overstates their emissions: disclosure lowers the recorded value of a firm below the estimate, whereas a firm above it gains from staying silent (Section 4, Figure 1). H2 therefore predicts that few switchers report more than the vendor had estimated, that switchers come mainly from the clean end of their industry, with a mean entry rank clearly below 0.5, and that switchers near the industry median show a gap close to zero. As a first step, we describe the gap. Only 32% (31%) of switchers report more Scope 1 (Scope 2) than the vendor had estimated (positive gap). The other 68% (69%) report less (negative gap). Across all switchers, the first reported Scope 1 (Scope 2) value lies on average 64% (56%) below the estimate the firm would otherwise have received. This alone does not show selection, because the same pattern would arise if the estimate were too high for all firms that do not report.

Notably, switchers come from all parts of their industry, not mainly from the clean end (Table 6, Figure 3a). The cleanest tenth of reporting firms contains 15% of switchers, compared with 10% expected without selection. 45% of switchers rank above the industry median. The mean entry rank is 0.46. Thus, the average switcher sits slightly below the middle of the reporting firms of its industry. If firms switched mainly because they were overestimated, few switchers would come from the upper half.

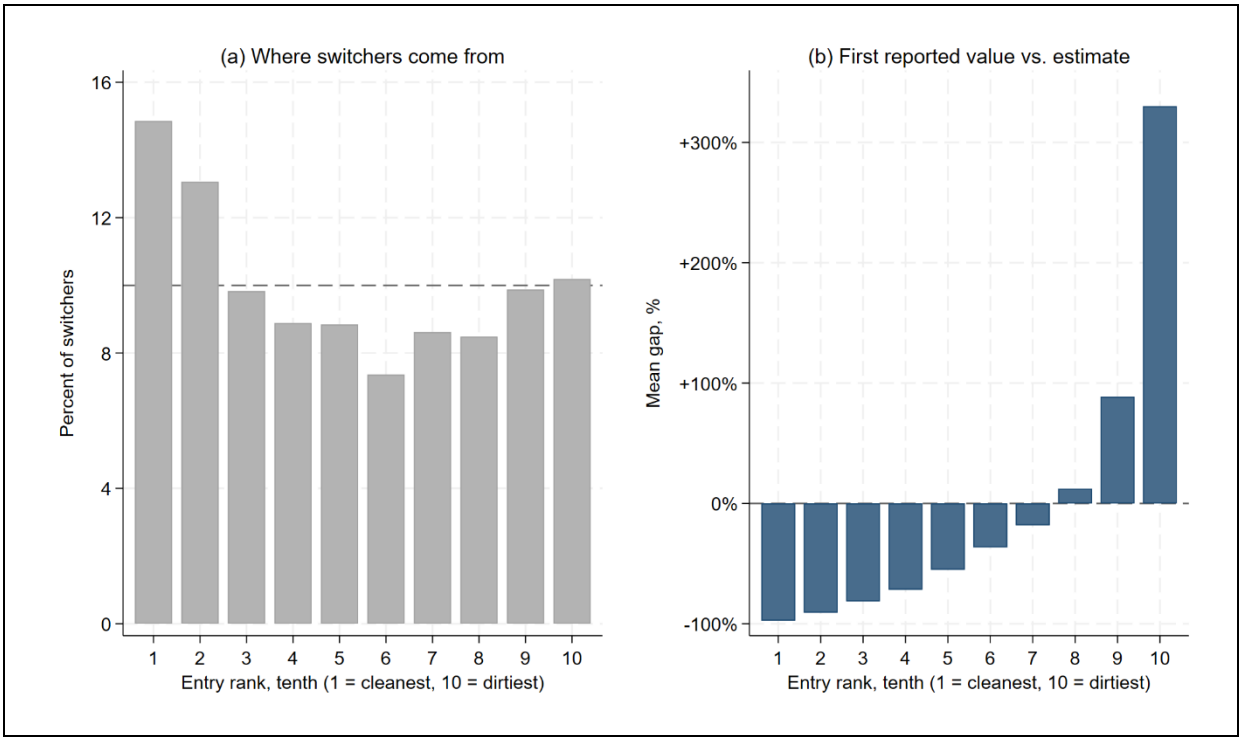
In addition, the gap is large not only for clean switchers (Table 6, Figure 3b). In the cleanest tenth, the first reported value lies on average 97% below the estimate. In the three dirtiest tenths switchers report on average more than the vendor had estimated. Under H2, such underestimated firms would have no reason to disclose. We therefore find little support for H2. Selection may add to the gap. Our tests compare switchers with firms that already report. If firms that never report are dirtier on average (Crippa et al., 2025; Chen et al., 2025), part of the gap of typical switchers could reflect that difference.

Table 6: Switchers by entry rank, mean gap a

Entry rank (tenth)	Share of switchers (%)	Mean gap (%)	Positive gap (%)
1 (cleanest)	14.9	−97	0.6
2	13.1	−91	0.7
3	9.8	−81	1.8
4	8.9	−72	11.7
5	8.8	−55	17.9
6	7.4	−37	32.5
7	8.6	−18	47.9
8	8.5	+12	65.1
9	9.9	+89	79.1
10 (dirtiest)	10.2	+330	92.0
No entry rank (fewer than ten reporting firms)	–	−58	33.2
All switchers	–	−64	32.4
Mean entry rank	0.463		
	(0.007)		
p-value: mean entry rank = 0.5	< 0.001		
Switchers	2,228	2,590	2,590

Entry rank (equation 3): position of the switcher's Scope 1 intensity (Scope 1 emissions per unit of revenue) among all firms that report Scope 1 in the same NAICS industry group and year, including the switcher; firms with estimated values are not included, values capped by LSEG are. Tenth 1 = cleanest, 10 = dirtiest; only industry-years with at least ten reporting firms. Share of switchers = share of the 2,228 switchers with an entry rank in each tenth; without selection, 10% each. Mean gap (%) = 100 × (exp(mean gap) − 1), with the gap from equation (2); negative = first reported value below the estimate the firm would otherwise have received. Positive gap = share of switchers that report more than estimated; about 50% if estimates were right on average and firms did not select. Mean entry rank: standard error clustered by industry in parentheses; p-value of a test that it equals 0.5. Switchers 2016–2023 with an industry classification.

Figure 3: Switchers by entry rank gap of switchers



Share of switchers in each tenth of the entry rank (equation 3); without selection, each tenth would contain 10% of switchers (dashed line). (b) Mean gap in each tenth, in % (100 × (exp(mean gap) − 1)); negative = first reported value below the estimate the firm would otherwise have received. Both panels show Table 6.

The underestimated switchers do not depend on this assumption. They disclosed although disclosure raised their recorded emissions, which is consistent with scrutiny of visible, high-impact firms (Marquis et al., 2016). Some switches may also reflect when LSEG starts to collect a firm's reports rather than when the firm decides to disclose (Section 2).

7. Robustness

Counting only reported values that LSEG accepts without capping, recorded emissions fall by 54% (Annex Table A4 and Figure A2). The drop is smaller because this definition leaves out the firms with the most extreme first values, most of which LSEG had raised to the industry's lower bound. With never-reporting firms as the only comparison group, the drop is 59%. Including firms whose value is later again estimated or extrapolated, and treating their switch as permanent, recorded emissions fall by 60%. A standard two-way fixed effects regression with controls for ln(total assets), return on assets and PP&E / total assets yields a smaller drop of 44%, as expected. Because the joint pre-trend test rejects in the main specification (Section 6.1), we also check how sensitive the drop is to deviations from parallel trends. Measuring all coefficients relative to the year before the switch shows a slight upward drift in the estimates of future switchers. Relative to comparison firms, they rise by about 0.17 log points from four years to one year before the switch (Annex Figure A3); extrapolating it would make the drop larger. The drop in the switch year remains significant unless deviations from parallel trends after the switch are more than four times as large as the largest deviation between two consecutive years before it (Rambachan and Roth, 2023; Annex Figure A4). Estimated separately by sector, the drop appears in eight of the ten sectors with at least 30 switches (Annex Table A5). Estimated separately by country, the drop is negative in all ten countries with at least 30 switches and significant in nine (Annex Table A6). The sector and country estimates use no control variables.

8. Conclusion

When a firm starts to report its Scope 1 and Scope 2 emissions, its recorded Scope 1 value in LSEG falls by about 60% and its Scope 2 value by about 51%. This drop is consistent with a change in measurement rather than a real reduction. It occurs exactly at the switch (Figure 2), drops of this size are rare when the data source does not change, and firms' operations do not shrink. Firms do not appear to switch mainly because they were overestimated; even a switcher

with typical emission intensity for its industry had been estimated at about twice its reported value.

Three recommendations follow for users of vendor data. First, a change in the method flag from estimated to reported should be treated as a change of measure, not as decarbonisation, for example when tracking portfolio intensity over time. Second, analyses that pool estimated and reported emissions should control for the method flag or use reported values only, because the two differ systematically in level. Third, event studies of disclosure should use estimators that are robust to staggered timing.

The results have limitations. We study one vendor and firms that start to report at some point; whether estimates are also too high for firms that never report cannot be observed directly. We observe the figures firms publish, not how they compute them, and reported values may themselves be understated (Cohen et al., 2026). Our results therefore show that the estimate lies above the firms' own figures, not necessarily above their true emissions. LSEG may also have recalculated past estimates with its current model. The switch year is the year in which LSEG first records the firm's own figure. It can lag the firm's first publication, and LSEG does not use first values from CDP that it considers extreme. Future research could compare vendors, use point-in-time data, and examine how vendors screen first disclosures and why the drop narrows in the years after the first disclosure.

Three questions that we leave out of this short paper deserve attention. First, why the estimate lies above the firms' own figures. LSEG's models are trained on reporting firms in the FTSE All-Cap index, and one of them derives intensities from economy-wide input-output tables (LSEG, 2025). Second, whether a high or rising estimate makes a firm start to report earlier. Third, whether firms that must report under a disclosure mandate show the same gap as voluntary reporters, which would separate the choice to report from the level of the estimate.

9. References

Acharya, V. v, DeMarzo, P., & Kremer, I. (2011). Endogenous Information Flows and the Clustering of Announcements. *American Economic Review*, *101*(7), 2955–2979. https://doi.org/10.1257/aer.101.7.2955

Aswani, J., Raghunandan, A., & Rajgopal, S. (2024). Are Carbon Emissions Associated with Stock Returns? *Review of Finance*, *28*(1), 75–106. https://doi.org/10.1093/rof/rfad013

Bolton, P., & Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of Financial Economics*, *142*(2), 517–549. https://doi.org/10.1016/j.jfineco.2021.05.008

Bolton, P., & Kacperczyk, M. (2026). Carbon disclosure and the cost of capital. European Accounting Review, 1–33. https://doi.org/10.1080/09638180.2026.2649612

Busch, T., Johnson, M., & Pioch, T. (2022). Corporate carbon performance data: Quo vadis? *Journal of Industrial Ecology*, *26*(1), 350–363. https://doi.org/10.1111/jiec.13008

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Chen, C. Y., Lioui, A., & Scaillet, O. (2025). *Green Silence: Double Machine Learning Carbon Emissions Under Sample Selection Bias*. https://doi.org/10.2139/ssrn.5368583

Cohen, L., Rouen, E., & Sachdeva, K. (2026). Widespread revisions of self-reported emissions by major US corporations. *Nature Climate Change*, *16*(1), 33–36. https://doi.org/10.1038/s41558-025-02494-9

Crippa, G., Berg, F., & Rigobon, R. (2025). *Corporate Omissions: Correcting the Bias in Carbon Reporting*. https://doi.org/10.2139/ssrn.5503418

Dye, R. A. (1985). Disclosure of Nonproprietary Information. *Journal of Accounting Research*, *23*(1), 123. https://doi.org/10.2307/2490910

El Ghoul, S., Gong, Z., & Guedhami, O. (2026). Understanding corporate emissions data: Implications for research design. Journal of Banking & Finance, 107822. https://doi.org/10.1016/j.jbankfin.2026.107822

Flammer, C., Toffel, M. W., & Viswanathan, K. (2021). Shareholder activism and firms' voluntary disclosure of climate change risks. *Strategic Management Journal*, *42*(10), 1850–1879. https://doi.org/10.1002/smj.3313

Frenkel, S., Guttman, I., & Kremer, I. (2020). The effect of exogenous information on voluntary disclosure and market quality. *Journal of Financial Economics*, *138*(1), 176–192. https://doi.org/10.1016/j.jfineco.2020.04.018

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277. https://doi.org/10.1016/j.jeconom.2021.03.014

Grossman, S. J. (1981). The Informational Role of Warranties and Private Disclosure about Product Quality. *The Journal of Law & Economics*, *24*(3), 461–483.

Haites, E., Bertoldi, P., König, M., Bataille, C., Creutzig, F., Dasgupta, D., de la rue du Can, S., Khennas, S., Kim, Y.-G., Nilsson, L. J., Roy, J., & Sari, A. (2024). Contribution of carbon pricing to meeting a mid-century net zero target. *Climate Policy*, *24*(1), 1–12. https://doi.org/10.1080/14693062.2023.2170312

Ilhan, E., Krueger, P., Sautner, Z., & Starks, L. T. (2023). Climate Risk Disclosure and Institutional Investors. *The Review of Financial Studies*, *36*(7), 2617–2650. https://doi.org/10.1093/rfs/hhad002

Kalesnik, V., Wilkens, M., & Zink, J. (2022). Do Corporate Carbon Emissions Data Enable Investors to Mitigate Climate Change? *The Journal of Portfolio Management*, *48*(10), 119–147. https://doi.org/10.3905/jpm.2022.1.410

Kim, E.-H., & Lyon, T. P. (2011). Strategic environmental disclosure: Evidence from the DOE's voluntary greenhouse gas registry. *Journal of Environmental Economics and Management*, *61*(3), 311–326. https://doi.org/10.1016/j.jeem.2010.11.001

König-Sykorova, M. (2026). Effectiveness of emission reduction policies and the role of corporate governance. *Energy Efficiency*, 19(1), 12. https://doi.org/10.1007/s12053-026-10414-9

Krueger, P., Sautner, Z., & Starks, L. T. (2020). The importance of climate risks for institutional investors. *The Review of Financial Studies*, *33*(3), 1067–1111. https://doi.org/10.1093/rfs/hhz137

LSEG (2025). LSEG greenhouse gas (GHG) emissions estimation model. Methodology, July 2025. London Stock Exchange Group.

Matsumura, E. M., Prakash, R., & Vera-Muñoz, S. C. (2014). Firm-Value Effects of Carbon Emissions and Carbon Disclosures. *The Accounting Review*, *89*(2), 695–724. https://doi.org/10.2308/accr-50629

Marquis, C., Toffel, M. W., & Zhou, Y. (2016). Scrutiny, Norms, and Selective Disclosure: A Global Study of Greenwashing. *Organization Science*, *27*(2), 483–504. https://doi.org/10.1287/orsc.2015.1039

Milgrom, P. R. (1981). Good News and Bad News: Representation Theorems and Applications. *The Bell Journal of Economics*, *12*(2), 380. https://doi.org/10.2307/3003562

Rambachan, A., & Roth, J. (2023). A More Credible Approach to Parallel Trends. *Review of Economic Studies*, *90*(5), 2555–2591. https://doi.org/10.1093/restud/rdad018

Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting and Economics*, *5*, 179–194. https://doi.org/10.1016/0165-4101(83)90011-3

Zamarioli, L. H., Pauw, P., König, M., and Chenet, H. (2021). The climate consistency goal and the transformation of global finance. *Nature Climate Change*, 11(7), 578–583. https://doi.org/10.1038/s41558-021-01083-w

Annex

A.1 Estimator

Let Y_{it} be ln(Scope 1 + 1) of firm i in year t, G the year in which a firm switches, and C_t the firms that have not switched by year t (not-yet and never-reporting firms). For each switch year g and year t, Callaway and Sant'Anna (2021) estimate the average effect of the switch on the switchers of year g,

$$ATT(g,t) = E[Y_t - Y_{g-1} \mid G = g] - E[Y_t - Y_{g-1} \mid C_t] \quad \text{(A1)}$$

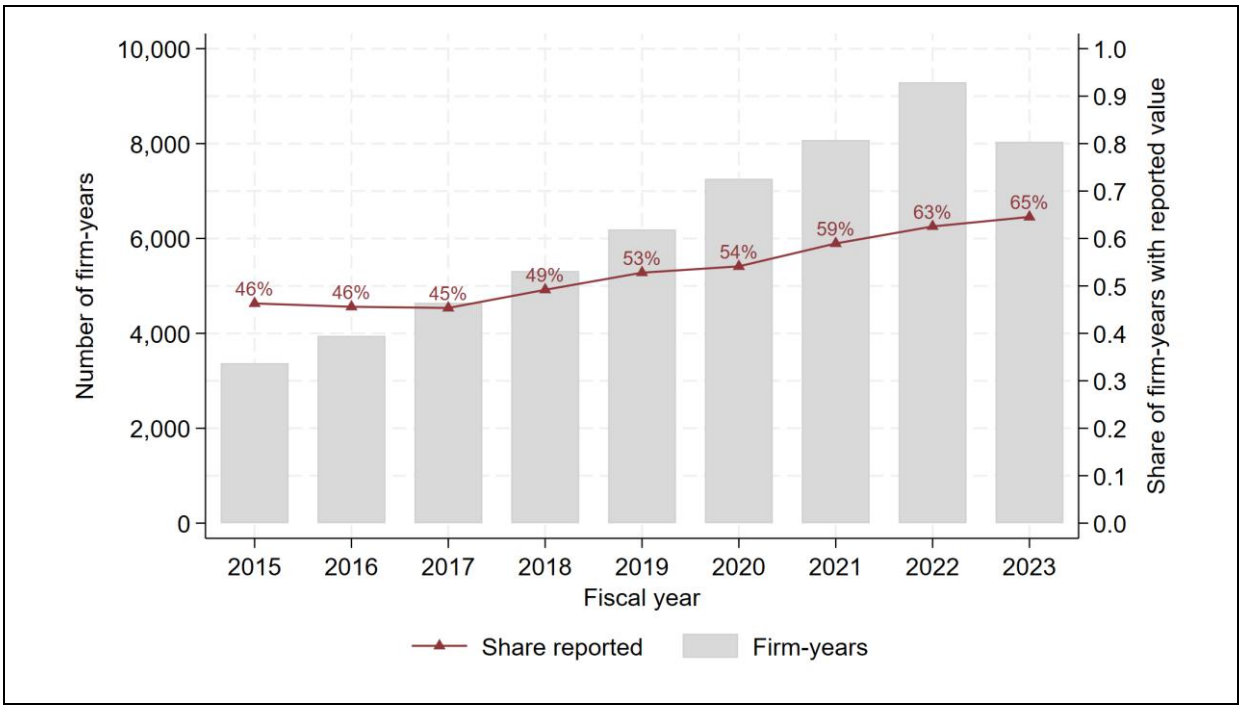
which equals δ_{gt} in equation (1). The event-study coefficient e years after the switch averages ATT(g, g + e) across switch years, weighted by the number of switchers. Pre-switch coefficients compare each year with the previous year. The main specification (Table 3, columns 3, 4, 7 and 8) includes the control variables $X_{i,g-1}$ of equation (1): ln(total assets), return on assets, PP&E / total assets and sector dummies, measured in the year before the switch. The estimator then compares switchers with comparison firms that had similar values: it estimates how the outcome of comparison firms changes with these variables and uses this to predict the change each switcher would have had without the switch (regression adjustment). Columns 1, 2, 5 and 6 of Table 3 use no control variables.

Table A1**.** Composition of the Scope 1 series and first disclosures by year

Year	Reported	Capped up	Capped down	Extra-polated	Energy model	Model estimate	Total	Share reported (%)	Switches (accepted)	Switches (first discl.)
2015	1,561	33	23	90	10	1,654	3,371	46.3	–	–
2016	1,800	35	22	80	10	2,000	3,947	45.6	156	171
2017	2,106	51	24	94	9	2,359	4,643	45.4	207	229
2018	2,613	67	36	95	10	2,492	5,313	49.2	249	283
2019	3,268	92	41	98	12	2,679	6,190	52.8	437	474
2020	3,927	124	48	140	11	3,006	7,256	54.1	348	399
2021	4,762	148	59	148	12	2,949	8,078	59.0	447	494
2022	5,812	167	72	164	18	3,061	9,294	62.5	449	487
2023	5,188	143	62	191	13	2,441	8,038	64.5	298	333
Total	**31,037**	**860**	**387**	**1,100**	**105**	**22,641**	**56,130**	**55.3**	**2,591**	**2,870**

Firm-years by LSEG method flag, 2015–2023, excluding Finance and Insurance. Capped up = Winsorized_low, capped down = Winsorized_high, extrapolated = Extrapolated and Energy_extrapolated, energy model = Energy_model, model estimate = Aggregated_model. Switches are first years with a reported value (accepted) or with a reported or capped value (first disclosure), counted only if the previous firm-year carries a model estimate.

Figure A1: Reporting status of Scope 1 emissions, 2015–2023



Source: own calculations based on LSEG data.

Table A2: Distribution of firms and switches by sector

	Firms	Disclosed (%)	Switches	Median gap (%)
Administrative, support and waste management	162	60.0	53	−79
Professional, scientific and technical services	747	40.0	152	−74
Information	892	42.8	241	−74
Construction	489	55.0	138	−74
Wholesale trade	319	57.8	108	−70
Retail trade	476	54.8	135	−68
Health care and social assistance	155	37.1	42	−66
Manufacturing	3,951	61.1	1,080	−50
Real estate, rental and leasing	624	55.4	226	−47
Utilities	376	72.3	87	−41
Mining, quarrying, oil and gas extraction	515	62.0	157	−23
Transportation and warehousing	375	69.9	123	+30
Accommodation and food services	168	60.7	46	+47
Other sectors (6)	255	43.9	62	–
Sector missing	563	66.1	220	–
Total	**10,067**	**57.5**	**2,870**	**−54**

Sample of 56,130 firm-years of 10,067 firms, 2015–2023, excluding Finance and Insurance. NAICS sectors, sorted by the median gap from the largest drop to the smallest. Disclosed = share of firm-years with a reported or capped value. Switches = first disclosures. Median gap = median of equation (2) across switchers, expressed as a percentage change, 100 × (exp(Gap) − 1); not computed for firms without an industry classification. Other sectors: Arts, Entertainment and Recreation; Agriculture, Forestry, Fishing and Hunting; Educational Services; Other Services; Management of Companies; Public Administration.

Table A3: Distribution of firms and switches by country of headquarters

	Firms	Disclosed (%)	Switches	Median gap (%)
Sweden	271	60.5	64	−85
Taiwan	216	82.9	53	−80

	Firms	Disclosed (%)	Switches	Median gap (%)
South Korea	175	70.2	46	−78
Thailand	199	77.7	56	−76
Singapore	105	60.4	45	−74
Japan	551	76.0	172	−73
Malaysia	559	51.1	100	−69
Switzerland	152	69.5	61	−68
United Kingdom	445	87.6	82	−63
China	1,278	38.8	453	−63
South Africa	101	77.2	19	−63
Hong Kong	168	71.8	85	−60
Australia	314	40.9	118	−56
United States	2,486	39.9	756	−48
India	604	72.5	122	−44
France	166	81.5	31	−31
Brazil	115	76.6	28	−24
Canada	331	54.9	102	−20
Germany	245	73.0	56	0
Italy	101	90.8	9	+1
Other countries	1,485	65.3	412	–
Total	**10,067**	**57.5**	**2,870**	**−54**

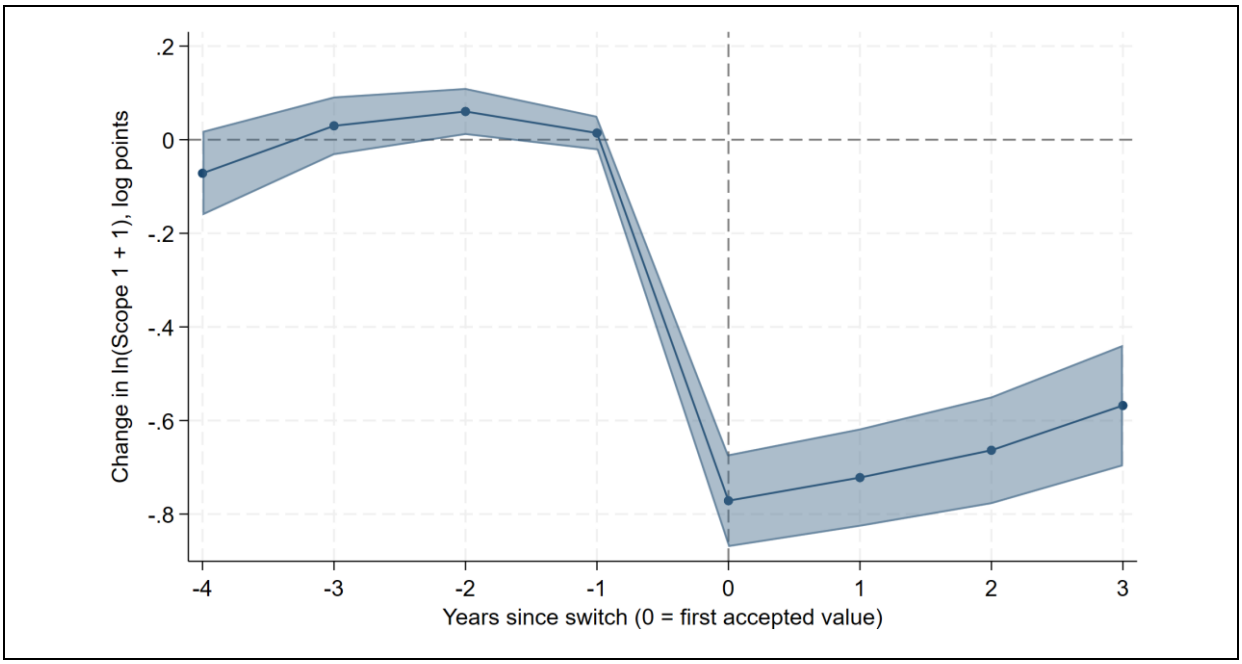
Same definitions as Table A2. The 20 countries with the most firm-years, sorted by the median gap from the largest drop to the smallest.

Table A4: Robustness of the drop in the switch year

Specification	Drop (log points)	SE	Change (%)	Comparison firms / controls
Baseline (Table 3, column 3)	−0.913***	(0.053)	−60	Not yet / never reporting; ln(total assets), ROA, PP&E / assets, sector
Only values accepted by LSEG without capping	−0.771***	(0.051)	−54	Not yet / never reporting; as row 1
Without control variables (Table 3, column 1)	−0.968***	(0.049)	−62	Not yet / never reporting; none
Never-reporting firms as comparison group	−0.898***	(0.053)	−59	Never reporting; as row 1
Two-way fixed effects, all firm-years	−0.584***	(0.030)	−44	Firm and year fixed effects; ln(total assets), ROA, PP&E / assets
Including firms whose value is later again estimated or extrapolated	−0.919***	(0.051)	−60	Not yet / never reporting; as row 1
Country dummies added to the control variables	−0.912***	(0.053)	−60	Not yet / never reporting; as row 1, plus country dummies

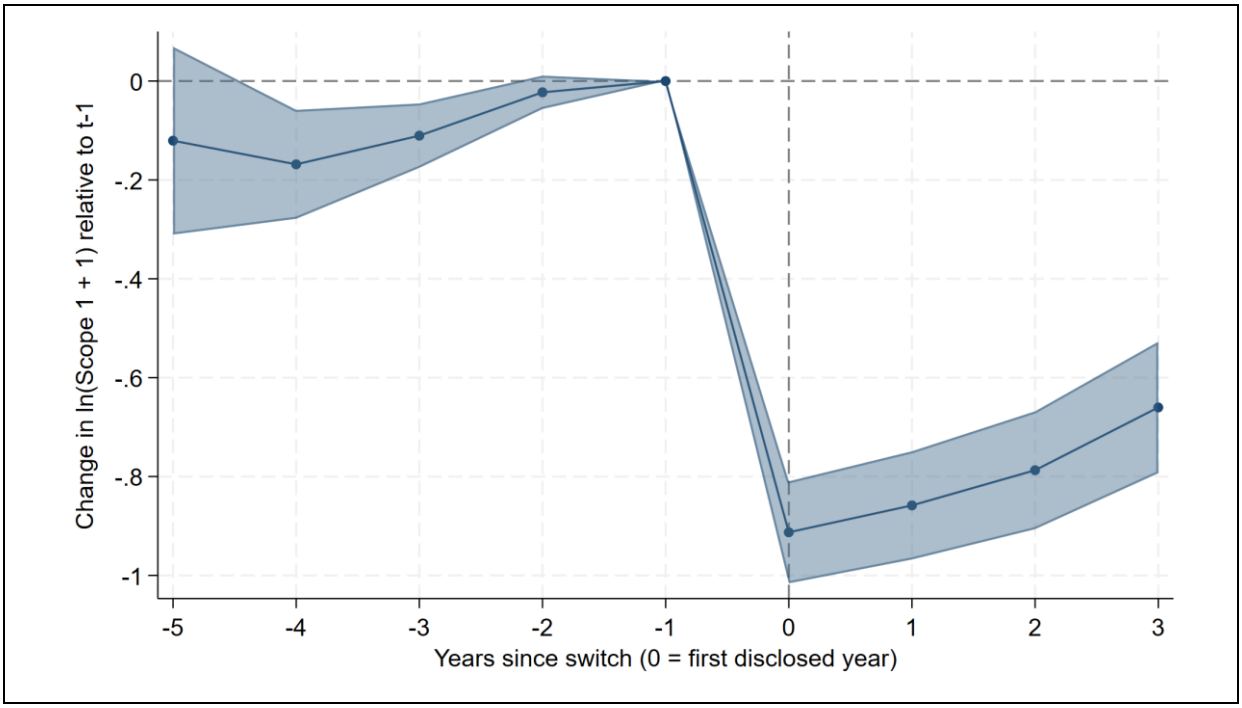
Drop = coefficient in the switch year, ln(Scope 1 + 1). Rows 1–4, 6 and 7: Callaway and Sant'Anna (2021), firms switching 2016–2020. Rows 1, 2, 4, 6 and 7 use the control variables of the main specification (equation 1), row 3 none. In row 2, firms whose own figures LSEG always caps are not used as comparison firms. Row 5: two-way fixed effects regression of ln(Scope 1 + 1) on an indicator for an accepted reported value, standard errors clustered by firm. Row 6: as row 1, but including firms whose value is later again estimated or extrapolated; their switch is treated as permanent. Row 7: as row 1, with dummies for the ten countries of headquarters with at least 30 switchers added to the control variables; firms from other countries form the base group. Change = 100 × (exp(drop) − 1). * p < 0.10, ** p < 0.05, *** p < 0.01.

Figure A2: Event study, only values accepted by LSEG without capping



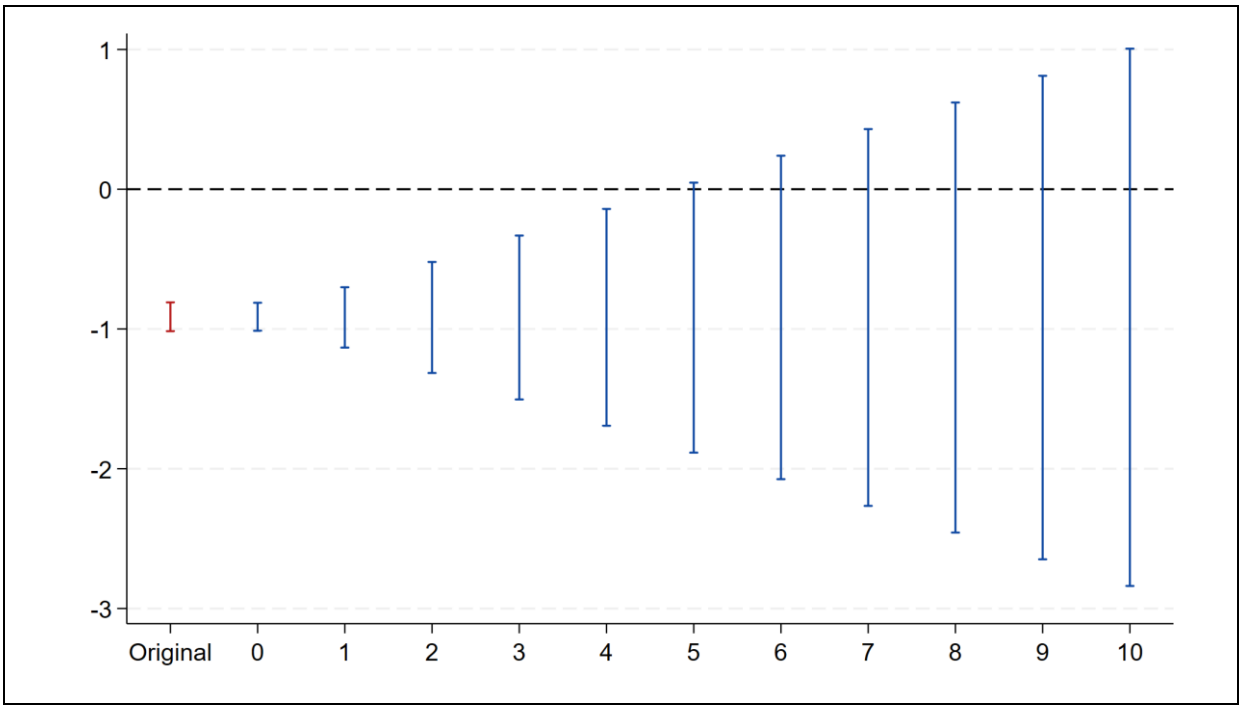
Callaway and Sant'Anna (2021) event-study estimates with 95% confidence intervals; see Table A4, row 2.

Figure A3: Event study with all coefficients relative to the year before the switch



Scope 1, main specification with control variables. All coefficients relative to t − 1.

Figure A4: Sensitivity of the drop in the switch year to violations of parallel trends



Robust 95% confidence intervals for the drop in the switch year following Rambachan and Roth (2023), relative magnitudes, based on the estimates in Figure A3. M (horizontal axis) = largest allowed deviation from parallel trends after the switch, as a multiple of the largest deviation between two consecutive years before the switch; Original = conventional confidence interval. The interval excludes zero up to M = 4.

Table A5: Recorded Scope 1 emissions at the switch, by sector

Sector	Switchers	ATT(0)		Change (%)	ATT(3)		Pre-trend p
Manufacturing	470	−0.943***	(0.074)	−61	−0.877***	(0.085)	0.052
Real estate, rental and leasing	123	−0.635***	(0.180)	−47	−0.516**	(0.211)	0.455
Mining, quarrying, oil and gas extraction	92	−0.228	(0.155)	−20	0.025	(0.251)	0.152
Information	83	−1.322***	(0.180)	−73	−0.969***	(0.198)	0.223
Retail trade	68	−1.172***	(0.180)	−69	−0.961***	(0.217)	0.103
Professional, scientific and technical services	65	−1.348***	(0.264)	−74	−1.383***	(0.262)	0.095
Construction	60	−1.166***	(0.247)	−69	−0.841***	(0.270)	0.198
Wholesale trade	60	−1.260***	(0.254)	−72	−0.762***	(0.260)	0.033
Transportation and warehousing	52	0.183	(0.214)	+20	0.208	(0.341)	0.038
Utilities	52	−1.243***	(0.316)	−71	−0.733*	(0.402)	0.230

Callaway and Sant'Anna (2021) estimates for ln(Scope 1 + 1), estimated separately for each NAICS sector with at least 30 switchers in the main sample (cohorts 2016–2020, first disclosure, reverters excluded). Controls are not-yet- and never-reporting firms of the same sector; no covariates. ATT(0) and ATT(3) = gap in the switch year and three years later, in log points; standard errors in parentheses. Change = 100 × (exp(ATT(0)) − 1). Pre-trend p = p-value of the joint test that all pre-switch coefficients are zero. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table A6: Recorded Scope 1 emissions at the switch, by country of headquarters

Country	Switchers	ATT(0)		Change (%)	ATT(3)		Pre-trend p
United States	427	−0.869***	(0.079)	−58	−0.859***	(0.092)	0.804
China	142	−1.201***	(0.160)	−70	−0.698***	(0.173)	0.005
Japan	113	−1.012***	(0.140)	−64	−0.676***	(0.205)	0.072
Hong Kong	61	−1.554***	(0.295)	−79	−2.324**	(1.073)	0.871
Canada	58	−0.645***	(0.245)	−48	−0.693**	(0.311)	0.683
Australia	45	−0.318	(0.314)	−27	−0.125	(0.370)	0.147
South Korea	35	−1.564***	(0.278)	−79	−1.190***	(0.390)	0.174
Taiwan	34	−1.744***	(0.360)	−83	−3.412***	(0.483)	–
United Kingdom	33	−1.050***	(0.348)	−65	−1.753***	(0.521)	–
Germany	30	−0.687***	(0.239)	−50	−0.193	(0.256)	0.243

As Table A5, estimated separately for each country of headquarters with at least 30 switchers in the main sample. Controls are not-yet- and never-reporting firms of the same country. – = pre-trend test not available because not all pre-switch years are observed.

Declaration on the use of generative AI

During the preparation of this work, the author used Claude (Anthropic), a generative AI assistant, for the following tasks: writing and revising the STATA code (do-file) for data preparation, estimation and figures; drafting, restructuring and language editing of the manuscript in English, including the incorporation of comments on earlier drafts; searching for and checking literature; suggesting additional analyses and robustness tests; and supporting to develop the schematic illustration in Figure 1. The author defined the research question, obtained the data, ran all analyses and decided on the content of the paper. The author reviewed and edited all AI-generated output, checked results and references against the original sources, and takes full responsibility for the content of this paper.