A Report on “Does Pricing Carbon Mitigate Climate Change? Firm-Level Evidence from the European Union Emissions Trading System” by Colmer et al. (2025)

Reviewer 2

February 11, 2026

v1

isitcredible.com

Disclaimer

This report was generated by large language models, overseen by a human editor. It represents the honest opinion of The Catalogue of Errors Ltd, but its accuracy should be verified by a qualified expert. Comments can be made here. Any errors in the report will be corrected in future revisions.

> I am wiser than this person; for it is likely that neither of us knows anything fine and good, but he thinks he knows something when he does not know it, whereas I, just as I do not know, do not think I know, either. I seem, then, to be wiser than him in this small way, at least: that what I do not know, I do not think I know, either.
>
> Plato, *The Apology of Socrates*, 21d

> To err is human. All human knowledge is fallible and therefore uncertain. It follows that we must distinguish sharply between truth and certainty. That to err is human means not only that we must constantly struggle against error, but also that, even when we have taken the greatest care, we cannot be completely certain that we have not made a mistake.
>
> Karl Popper, 'Knowledge and the Shaping of Reality'

Overview

Citation: Colmer, J., Martin, R., Muûls, M., and Wagner, U. J. (2025). Does Pricing Carbon Mitigate Climate Change? Firm-Level Evidence from the European Union Emissions Trading System. *Review of Economic Studies*, Vol. 92, No. 3, pp. 1625–1660.

Abstract Summary: The study estimates the efficacy of the European Union Emissions Trading System (EU ETS) using administrative data on regulated manufacturing firms. The EU ETS induced regulated manufacturing firms to reduce carbon dioxide emissions by 14–16% with no detectable contractions in economic activity, suggesting that the policy induced global emissions reductions through targeted investments in cleaner production processes.

Key Methodology: Matched difference-in-differences research design using comprehensive administrative data from the French manufacturing sector.

Research Question: Does pricing carbon mitigate climate change, and what is the firm-level evidence from the European Union Emissions Trading System?

Summary

Is It Credible?

Colmer et al. present a retrospective analysis of the European Union Emissions Trading System (EU ETS), utilizing administrative data from French manufacturing firms to argue that the policy successfully reduced carbon dioxide emissions by 14–16% without causing economic contraction. The authors position this as the "first evidence" supporting the efficacy of the EU ETS in delivering on its stated policy objective while avoiding carbon leakage (p. 1629). They propose that these results are driven by firms making fixed-cost investments in cleaner technologies, a mechanism they attempt to rationalize through an augmented theoretical model. While the study utilizes a sophisticated matched difference-in-differences design, the robustness of the headline claims is undermined by statistical fragility and unexamined alternative explanations.

The central claim of a 14–16% reduction in emissions appears fragile when subjected to rigorous robustness checks. The authors acknowledge that their primary matching specification leaves treated firms significantly larger and more emissions-intensive than their control counterparts (p. 1633). When the authors employ cardinality matching—a technique designed to minimize these covariate imbalances—the estimated reduction in Phase II effectively vanishes, dropping from -16.3% to a statistically insignificant -0.8% (p. 50). While this specification reduces the sample size by nearly half (to 85 treated firms), which naturally inflates standard errors, the disappearance of the effect in the most balanced comparison group suggests that the headline magnitude may be driven by residual differences between the treated giants and smaller control firms, rather than the policy itself. Consequently, the precise quantum of emissions reduction is likely less certain than the abstract suggests.

Furthermore, the authors' explanation for the "no economic contraction" result—

that firms adopted productivity-enhancing technology—lacks strong empirical support within their own analysis. The theoretical model introduced to explain the results relies on the assumption that new, cleaner technologies increase Total Factor Productivity (TFP), thereby offsetting regulatory costs (p. 1652). However, the empirical analysis finds no statistically significant effect on measured TFP (p. 1646). While the authors correctly note that a null result is preferable to the negative productivity shock predicted by standard models, it creates a disconnect between the theoretical rationalization and the data; the "win-win" outcome of reduced emissions and stable economic output remains an empirical observation in search of a confirmed mechanism. A plausible alternative explanation is the role of free permit allocations, which acted as a substantial subsidy during the study period. The authors dismiss this factor based on the lack of variation in allocation rules (p. 1631), but in the presence of real-world financial constraints, these transfers could explain why firms did not contract, rendering the "cost-free" nature of the regulation an artifact of compensation rather than innovation.

Finally, the extrapolation that the EU ETS induced "global emissions reductions" (p. 1625) is heavily conditioned by the specific context of the study. The analysis is limited to France, where the electricity grid was already 79% carbon-neutral during the study period (p. 1647). This unique environment constrained firms to abatement strategies focused on process efficiency rather than fuel switching, which might be the primary margin of adjustment in fossil-fuel-dependent economies. Consequently, the absence of leakage and the specific investment behaviors observed in France may not generalize to other jurisdictions within the EU ETS or globally, where the economic pressures and abatement options differ fundamentally.

The Bottom Line

The claim that the EU ETS reduced emissions in French manufacturing without economic cost is plausible but likely overstated in magnitude and certainty. The headline 14–16% reduction is highly sensitive to statistical specification, vanishing under the strictest matching conditions, and the mechanism attributed to productivity gains is not supported by the study's own TFP estimates. While the article effectively rules out catastrophic economic losses for the studied firms, this stability may be due to unexamined subsidies (free permits) and the specific advantages of France's low-carbon energy grid, limiting the result's applicability as a universal proof of cost-free climate policy.

Potential Issues

Fragility of the main emissions reduction finding: The headline claim of a 14–16% emissions reduction is sensitive to the choice of matching specification. While the main finding is robust to several alternative specifications, it becomes statistically insignificant when the authors employ cardinality matching—an algorithm designed to achieve better covariate balance. Under this specification, the estimated effect in Phase II disappears almost entirely, with the coefficient dropping to -0.008 from a baseline of -0.163 (Table B.6, p. 50). It is important to note that this specification reduces the sample size by nearly half, which increases standard errors and reduces statistical power (p. 1643). However, the fact that the core finding does not hold under the specification designed to produce the best-balanced control group raises questions about the robustness of the article's central conclusion.

Unbalanced matched sample: The study's difference-in-differences identification strategy relies on the parallel trends assumption, which is argued to be more credible after matching regulated (treated) firms with similar unregulated (control) firms. However, the matched groups remain profoundly different on key baseline characteristics. According to the article's own data for the baseline year 2000, the average treated firm was still vastly larger and more polluting than the average control firm, even after matching. Specifically, treated firms emitted 2.57 times more CO_2, were 1.56 times more capitalized, and were approximately 2.15 times more emissions-intensive than their matched counterparts, with all these differences being statistically significant at the 1% level (Table 1, p. 1633). The authors acknowledge that "some baseline differences remain" but argue they are "substantially smaller" than pre-match differences (p. 1634). While the difference-in-differences design controls for time-invariant level differences and the authors provide visual evidence of parallel pre-trends (Figure 2, p. 1636), the large, systematic differences in scale and pollution intensity raise questions about whether these fundamentally different types of firms would have followed the same economic trajectory in the absence of the policy, particularly through a major shock like the Great Recession.

Inconclusive empirical support for the theoretical model's key mechanism: The article proposes a theoretical model of "technology switching" to explain how firms reduced emissions without economic contraction. A key feature of this model is that the new, cleaner technology also increases Total Factor Productivity (TFP), offsetting the costs of carbon pricing. However, the article's own empirical analysis does not find statistically significant evidence for this TFP increase. The authors report that "the EU ETS has no effect on measured TFPR" (p. 1651), with the point estimate being statistically indistinguishable from zero (Table 3, p. 1646). The authors argue that the positive (though insignificant) point estimates for TFP, value added, and employment are directionally consistent with their model and inconsistent with a standard model that predicts economic contraction (p. 1651). While the results may rule out the alternative hypothesis of contraction, the lack of robust statistical evidence for the TFP increase means the central mechanism of their explanatory model is not empirically confirmed.

Weak causal evidence for the specific investment mechanism: The article's primary explanation for the emissions reductions is that the EU ETS induced firms to make fixed-cost investments in "integrated production technologies." The direct evidence for this specific mechanism is drawn from survey data that began in 2001, one year after the study's main baseline period. As the authors acknowledge, this data limitation means "we are unable to investigate whether trends in those outcomes are parallel during the pre-announcement period" (p. 1647). The absence of a pre-treatment period for the specific pollution-abatement investment variables undermines the difference-in-differences design for this particular outcome. While the article does provide more robust evidence for an increase in the general capital stock, for which pre-trends are shown to be parallel (Table 2, p. 1641; Figure 3e, p. 1642), the evidentiary chain linking the policy specifically to pollution-abatement investment is weaker.

Potential overstatement of global impact: The article concludes that the EU ETS "induced global emissions reductions" (p. 1625). This strong claim is an inference based on finding significant local emissions reductions in France combined with no evidence of carbon leakage through the specific channels the study could measure, such as increased imports or firm contraction. However, the study's design, which is confined to French manufacturing firms, cannot rule out other significant forms of leakage, such as a shift in production from a regulated French firm to an unregulated firm in another country, or a shift in demand from a French ETS firm's product to that of an unregulated foreign competitor. The authors' conclusion conflates the absence of evidence for specific, measurable leakage channels with evidence of absence for all leakage channels, which is a debatable inferential leap. The evidence more directly supports a narrower conclusion about emissions reductions at French manufacturing firms with no detected local leakage.

Limited external validity due to France's unique energy context: The study's findings are based exclusively on firms in France, a country with a unique, largely carbon-neutral electricity grid. The authors note that "79% of the electricity generated in France in 2012 was carbon neutral" (p. 1647). This context is a significant constraint on the generalizability of the results. In countries with fossil-fuel-based electricity grids, the costs and benefits of abatement strategies like electrification are fundamentally different. The article finds that firms primarily invested in process efficiency rather than switching to electricity (p. 1647). While this finding makes the cleanliness of the grid less directly relevant to the specific mechanism observed, it also means the study offers little insight into how firms might behave in jurisdictions where fuel switching is a more viable or economically attractive abatement option. This limits the direct applicability of the article's conclusions about abatement mechanisms to other countries with different energy systems.

Unexamined role of free permit allocation: The article's central finding is that regulated firms reduced emissions by making capital investments without suffering a contraction in economic activity. The analysis attributes this to firms adopting new, more productive technologies. However, the article does not empirically examine a significant alternative explanation: the free allocation of emissions permits. During the study period, this was the primary allocation method and represented a substantial financial transfer to regulated firms. These transfers could have directly financed the observed investments, particularly for credit-constrained firms, thereby explaining the lack of economic harm. The authors cite standard economic theory that "permit allocation should not affect firm behaviour at the margin" and state they "lack a credible strategy to test for a causal effect" (p. 1631). While theoretically justified in a frictionless model, dismissing the potential income effect of a large asset transfer without an empirical test for the role of financial constraints leaves a primary, real-world explanation for the article's key economic finding unexamined.

Measurement of emissions: The study's measure of CO2 emissions is derived from fuel consumption data and therefore primarily captures combustion-related emissions, excluding process emissions that result from chemical reactions in industries like cement manufacturing. The authors note in the supplementary material that process emissions constitute a significant portion of the total, estimating that "a quarter of GHG emissions from manufacturing are process emissions" in France (p. 39). However, they perform a validation exercise comparing their combustion-based measure to total verified emissions from the official EU registry (which includes process emissions), finding a correlation of 0.96 (p. 39). This high correlation suggests that the combustion data is a very strong proxy for total emissions for the vast majority of firms in the sample. Nonetheless, by construction, the measure may not fully capture abatement dynamics for the subset of firms where process emissions are dominant.

Methodological limitations in secondary analyses: The study contains several acknowledged methodological limitations in its secondary and robustness analyses that, taken together, introduce uncertainty. First, the analysis of firm productivity (TFPR) relies on a "simplified version" of estimation that uses OLS and may not adequately address the endogeneity of input choices, a known source of bias in productivity measurement (p. 1651). Second, the measurement of carbon leakage through imports is based on coarse, sector-level proxies for carbon intensity and yields statistically imprecise estimates, making it difficult to rule out a substantial increase in imports (p. 1645). Third, the matching procedure includes the baseline level of CO2 emissions as a covariate, a practice that can introduce bias through regression to the mean; the authors mitigate this concern by demonstrating parallel pre-trends over several years and matching on a vector of other firm characteristics (p. 1634).

Future Research

Investigation of free allocation effects: Future research should explicitly test the role of free permit allocations in preventing economic contraction. By exploiting variation in allocation levels relative to emissions across firms or sectors, researchers could determine whether the observed stability in value added and employment is a result of technological innovation or simply a wealth effect from the transfer of valuable assets.

Reconciliation of TFP and investment: Future work is needed to resolve the discrepancy between observed capital investments and stagnant Total Factor Productivity. Research could focus on distinguishing between "productive" capital and "abatement" capital to understand why the hypothesized efficiency gains from new technologies did not materialize in the productivity data, perhaps by using more granular data on specific machinery purchases.

External validity in carbon-intensive grids: Future studies should replicate this matched difference-in-differences design in EU ETS countries with high carbon intensity in their electricity grids, such as Poland or Germany. This would test whether the "no leakage" and "process innovation" findings hold when fuel switching is a viable option and when the pass-through costs of electricity are significantly higher.

isitcredible.com