Adopt or Innovate: Understanding technological responses to cap-and-trade Online Appendix

Raphael Calel*

September 17, 2019

I. Defining low-carbon patents

Low-carbon innovation has long played an important role in national climate policy debates as well as international climate change negotiations, but individual researchers had struggled to settle on common metrics to measure its scale, distribution, and progress. Starting in 2009, the European Patent Office, United Nations Environment Programme, and the International Centre for Trade and Sustainable Development jointly undertook to create a new patent class covering technologies that control, reduce, or prevent greenhouse gas emissions, with the hope that this would enable more research on low-carbon innovation and inform public policy.

The European Patent Office had patent examiners specialised in the relevant technologies, supported by external experts, conduct a series of patent searches—looking at European and International patent classification codes, at patent abstracts, and even at the text of the claims. These searches were used to populate a database with low-carbon patents, and their search strategies were codified and refined over several iterations until they yielded an automated search algorithm capable of producing reliable results. The new patent class, labeled "Y02," was unveiled in 2010, and this algorithm is now used to automatically identify all new low-carbon technologies added to the database. The algorithm

*Calel: McCourt School of Public Policy, Georgetown University, and Grantham Research Institute on Climate Change and the Environment, London School of Economics. Email: raphael.calel@georgetown.edu.

is maintained and updated by patent examiners with relevant technological expertise, and any significant changes in the algorithm are applied retrospectively as well as prospectively, to make available a consistent time-series to researchers.

Patents were originally identified for two subclasses of technologies: capture, storage, sequestration, or disposal of greenhouse gases (Y02C), and reduction of greenhouse gas emissions related to energy generation, transmission, or distribution (Y02E). Over time, the class has expanded to cover a broader range of technologies, including some adaptation technologies (Y02A) as well as some energy efficiency technologies (in Y02B and Y02D especially, and to a lesser degree in Y02P). Table 1 provides descriptions of the current list of technology subclasses.

The EU ETS caps direct emissions, but does not regulate adaptation, or indirect emissions resulting from electricity consumption. As elaborated in the paper, I would expect the EU ETS to give capped firms greater incentives to develop technologies that reduce emissions directly, compared with uncapped firms. But the same cannot be said of all the technologies represented in the Y02-class. The EU ETS is not likely to have much effect on technologies relating to adaptation, for instance. The EU ETS could well encourage energy efficiency innovation if it raised electricity prices, but that would likely affect capped and uncapped firms in equal measure.

These conceptual categories are a bit fuzzy in practice. Adaptation technologies are perhaps the easiest to separate out (Y02A), so I do not to count them as "low-carbon patents" in my analysis. My focus is on mitigation technologies. But some of the other subclasses are trickier, since they appear to include some mixture of emissions saving technologies and energy saving technologies. Trickier still, about 15% of patents are tagged in two or more subclasses, and even technologies tagged only once can have different effects depending on their application, e.g. new building insulation materials reduce emissions in buildings heated by traditional boilers, but save electricity for electrically heated buildings. Clearly, no empirical definition of "low-carbon patents" can be made to align precisely with the conceptual distinctions.

To gauge how serious this misalignment is, the last column of table 1 shows the distribution of patents across subclasses, counting only UK patents filed by British firms from 2000 to 2012. I count the number of patents tagged in each subclass and divide patents with multiple tags equally across subclasses to avoid double counting. When I do this, I see that the subclasses most likely to contain energy efficiency technologies (Y02B and Y02D) make up 16% of Y02 patents.

Table 1: Technologies included in the Y02 patent class

Subclass	Description	Technology examples	Share
Y02A	Adaptation to climate change	Flood resilient electrical equipment, Plants tolerant of drought salinity or heat, Water efficient irrigation, Early warning systems for extreme weather events.	7.0%
Y02B	Buildings	Integration of renewable energy sources in buildings, Fuel efficient boilers, Waste heat powered water heating, Combined cooling heat and power generation (trigeneration), Energy efficient lighting, Insulation.	9.9%
Y02C	Capture and storage of greenhouse gases	CO_2 capture, Subterranean CO_2 storage, N_2O disposal, Methane capture.	15.5%
Y02D	Information and communication technology	Energy efficient computing, Energy-aware routing, Power-based selection of communication route or path in wireless communication networks.	6.5%
Y02E	Production, distribution and transport of energy	Renewable energy generation, Biofuel production, Nuclear power, Combined heat and power generation, Combined cycle power, Superconducting power lines, Fuel cells, Batteries.	30.2%
Y02P	Industry and agriculture	Recycling CO_2-rich gas in metals processing, Catalytic reduction of N_2O emissions from adipic acid and caprolactam production (chemicals used in plastics and nylon), Cements with lower clinker content, CO_2 capture for large oxy-fuel furnaces used in glass production, High-efficiency and renewable fuel powered ceramics kilns, Agricultural methane capture, Plants with high carbon sequestration potential, Inventory and reporting systems for greenhouse gases.	15.2%
Y02T	Transportation	High-efficiency internal combustion engines, Hybrid vehicles, Batteries for electric vehicles, Charging systems for electric vehicles, Weight and drag reduction technologies for airplanes, Hydrodynamically efficient hull and propeller designs.	23.7%
Y02W	Waste and wastewater	Biogas capture and recycling, Landfill sealing and gas capture, Production of fertilisers from organic waste, Recycling of batteries.	5.9%

To get a better sense of the share of energy efficiency patents, I have attempted to identify patent sub-sub-classes (down to the 10-digit level) that cover technologies that are electrically powered or where the main application is to reduce electricity use. At the most disaggregated level, it is possible to distinguish heating systems that use heat pumps (typically electrically powered), as opposed to those using condensing boilers (typically gas-fired or oil-fired). I can isolate patents for induction furnaces used in metals processing, for microwave oven-based industrial food processing techniques, and for other electrically powered technologies across all industries.

I find that 15% of Y02 patents are tagged in energy efficiency sub-sub-classes, and 12% are tagged in energy efficiency sub-sub-classes but *not* in other Y02 subclasses. These 12% of patents are the ones most likely to only have energy efficiency benefits, and might therefore provide a reasonable estimate of how inflated the counts of "low-carbon patents" might be relative to the conceptual definition. Still, since there are likely to be some technologies with potential to reduce direct emissions even within these sub-sub-classes, this number probably overstates problem.

Now that we have a sense of the magnitude of the discrepancy, it's worth taking a moment to think through the consequences. Suppose the EU ETS disproportionately encourages emissions saving patenting for capped firms, but encourages energy efficiency patenting in equal measure for capped and uncapped firms. The difference in the absolute number of "low-carbon patents" filed by suitably matched capped and uncapped firms will only capture additional emissions saving patenting. Since energy efficiency patents appear on both sides of the ledger, they will not affect the difference between them. Even when the patent counts include some number of energy efficiency technologies, then, the difference in the number of low-carbon patents will measure the relative effect of the EU ETS on emissions saving patenting. The EU ETS may affect energy efficiency patenting as well, but this effect would be added on top of my estimate.

The share of energy efficiency technologies matters more when I try to express the number of additional low-carbon patents in proportional terms, rather than in absolute numbers. As we've just seen, the numerator only includes the additional emissions saving patents. Ideally, my count of "low-carbon patents" in the denominator would include only emissions saving technologies as well. The more energy efficiency patents that are counted in the denominator, the smaller the proportional effect will appear. Because of the unavoidable fuzziness of the categories, the proportional effect will tend to be an underestimate true proportional effect on emissions saving technologies.

To illustrate the magnitude, suppose the true effect of the EU ETS was to add 1 extra emissions saving patent for every 4 filed in the counterfactual scenario, a 25% increase. But if the denominator was inflated by 12%, to take my earlier approximation, my estimated effect would be 22.3% instead of 25%. If the denominator was inflated by 25%, my estimated effect would be 20% instead of 25%. This gives us a rough idea of how seriously I may underestimate the proportional effect of the EU ETS on emissions saving patents.

The difficulty of categorising the technologies could also influence the estimated effect on all other non-low-carbon patents. Whenever the search algorithm fails to tag an emissions saving patent, that patent ends up in the "other" category. Part of the EU ETS's effect on emissions saving technologies might then be lumped in with the effect on "other" patents. The difference between capped and uncapped firms in the total number of patents filed would accurately estimate the EU ETS's relative effect on patenting, but the breakdown between "low-carbon" and "other" would tend to understate the contribution made by low-carbon patents.[1] This is unlikely to have much influence in practice, though, since misclassified emissions saving technologies are going to be such a small proportion of "all other patents."

In sum, although the Y02-class is largely made up of patents that protect new emissions savings technologies, it does not (and cannot be made to) map onto this conceptual category exactly. Importantly, though, the difference in the numbers of Y02-tagged "low-carbon patents" filed by suitably matched capped and uncapped firms will capture the EU ETS's effect on emissions saving technologies, and omit its effect on energy efficiency and other technologies. The proportional effect will tend to be overly conservative, which is a reason for focusing attention on the absolute effect, even if it is sometimes harder to get a sense of its magnitude. To the extent that the Y02-class fails to tag some emissions saving patents, the estimate of the absolute effect will be conservative, too, and the estimate of the program's effect on "other" patents will be ever so slightly larger.

The difficulty of categorising technologies therefore leads to a number of potential biases, but they all point in the same direction: they make it less likely that I will find evidence of directed technological change favouring low-carbon innovation. My comparisons between capped and uncapped firms should therefore be interpreted as providing conservative estimates of the EU ETS's effect on low-carbon patenting.

II. Multi-plant companies

The EU ETS regulates all plants that exceed a certain threshold production capacity. But a company that operates an EU ETS plant may operate some smaller plants as well, so the

[1]The same holds for the difference-in-differences estimator. And both the simple- and double-difference would be preferable to a triple-difference estimator in this context. The triple-difference would pick up the effect on "other" patents, which may be due to misclassified emissions saving technologies, and end up with a lower estimate instead of a higher one.

share of regulated activities might differ across ETS firms. An ideal research design would distinguish companies along a continuum of treatment shares and leverage the variation in shares to study what sort of dose-response relationship characterises companies' adoption and innovation decisions. The presence of such a dose-response relationship would provide further evidence that the observed differences in outcomes are indeed an effect of the EU ETS.

I am unfortunately not in a position to perform this kind of analysis since the data used for this study are collected at the company-level and are not broken down to the plant-level. The only plant-level information I have used is the list of plants covered under the EU ETS, which I have linked up to the firm-level data sets in order to identify ETS companies (i.e. those operating at least one EU ETS plant). But these databases do not provide comparable data on plants below the regulatory thresholds.

To learn more about the how regulated shares might vary across EU ETS companies, we must search for other sources of plant-level information. For this purpose, I have managed to obtain and clean a historical version of the UK's Pollution Inventory. This database records annual CO_2 emissions from over 4,000 plants going back to 1998. It covers plants inside the Environment Agency's jurisdiction, which includes England but excludes Wales, Scotland, and Northern Ireland. It therefore does not cover all ETS plants and companies. Yet it is the only database I've found that provides consistent measurements of emissions across a large sample of ETS and non-ETS plants.

Of the 445 EU ETS companies in Britain, the Pollution Inventory includes data on plants belonging to 205 of them. They collectively operated 947 plants listed in the Pollution Inventory, of which 411 (or 43%) fell under the EU ETS sometime between 2005 and 2012. More than half are single-plant firms (figure 1, left panel), so the most common regulated share, measured by the number of plants, is obviously 100%. But the modal share is 100% even among just the multi-plant firms. For 73% of the firms in this sample, the EU ETS covers all of their plants. The distribution is heavily skewed by just two water utilities that operated a large number of plants, so even though only 43% of plants are regulated in the aggregate, the average firm has a regulated share of 85%.

But recall that the EU ETS is designed to cover the largest plants. A simple count of the number of regulated and unregulated plants will therefore systematically underestimate the share of regulated activity, and hence overstate the variation across firms. To address this, I compare the CO_2 emissions from each company's ETS plants and non-ETS plants (figure 1, right panel). Overall, the ETS plants accounted for 92% of reported CO_2 emissions in

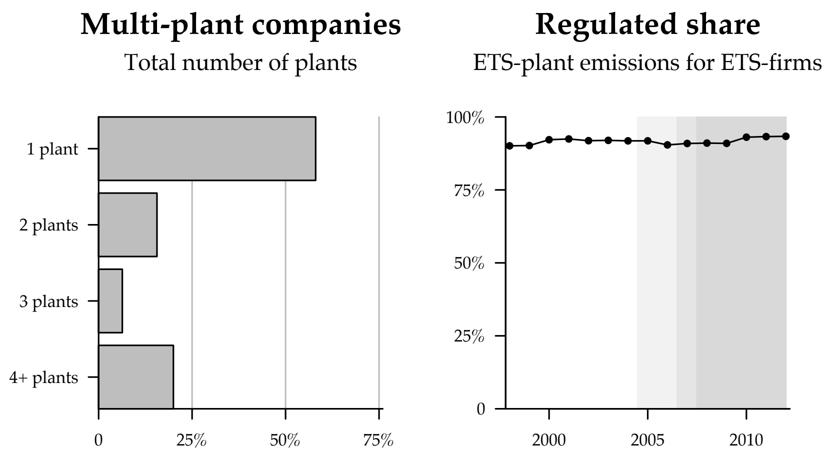

Figure 1: Multi-plant companies.
Notes: **Left panel:** The distribution of single-plant and multi-plant companies. **Right panel:** The share of CO_2 emissions from ETS-plants over time. Some plants are occasionally recorded as having emissions 'below reporting threshold,' which was 10,000 tonnes until 2011, and then raised to 100,000 tonnes. I have calculated upper and lower bounds based on imputing the emissions for these plants to either zero or the threshold value, but the figures differ by only a fraction of a percentage point. The line is plotted for the mid-point of this range. Both plots are based on a sample of 205 EU ETS companies whose plants are covered by the UK's Pollution Inventory.

the years leading up to the EU ETS, and 90% of the emissions from multi-plant firms. The distribution consists of a very large number of observations at or near 100%, along with a small number of low-lying outliers. As seen in figure 1, these shares remain stable even after the EU ETS goes into effect, providing suggestive evidence that multi-plant companies are not re-balancing production towards their unregulated plants.

Although the EU ETS can in theory cover any share of a company's emissions, I find that there isn't much variation in practice. Based on data available for nearly half of British ETS companies, it seems as though the EU ETS covers all or nearly all of their emissions. This lack of variation among ETS companies implies (1) that a binary treatment indicator provides a highly accurate approximation of the underlying continuous treatment variable, and (2) that there does not exist sufficient statistical power to reliably estimate heterogeneous treatment effects with respect to the share of regulated activities.

Using a binary approximation of the treatment variable does somewhat restrict one's ability to use my estimates for prospective policy analyses. Conceptually, the difference between ETS and non-ETS companies measures the average treatment effect on the treated conditional on the distribution of regulated shares. These estimates are valid for a retrospective policy evaluation without modification, since the distribution of treatments has been fixed by the actual historical policy. But my estimates would not necessarily extrapolate in a straightforward manner if one wanted to forecast the consequences of expansions of the EU ETS. Such forecasts will also depends on how the distribution of regulated shares differs from the historical sample I used. If such an expansion mainly affected single plants that only accounted for small shares of the their operators' emissions, the results might well be quite different from my findings here.

The most important limitation of these estimates is that the Pollution Inventory does not necessarily contain the universe of plants operated by these 205 ETS companies. To the extent that they operate additional plants in Scotland, Wales, Northern Ireland, or elsewhere in Europe, my estimates might be biased downward (if those additional plants are predominantly in the EU ETS) or upward (if they are predominantly non-ETS). The estimates will be biased upward to the extent that they also operate plants outside of the 31 countries that participate in the EU ETS.

III. Matching design

Matching does not provide a basis for causal identification in and of itself, any more than least-squares regression. Matching is fundamentally a method for trimming and re-weighting observations so that the treated and control units appear more similar along a few select dimensions. But without a well-formulated research design, matching could just as well end up undermining causal identification by exacerbating differences along unobserved dimensions.

This study uses matching as a tool to isolate an underlying policy experiment. To this end, it is useful to spell out what the analogous 'true' experiment would look like, and how it relates to the EU ETS.

I start observing firms in the year 2000. At that time, every firm was a 'potential ETS firm.' Then in January of 2005, imagine I was asked to choose 272 firms to enter the EU ETS. But instead of just picking 272 firms, I selected 272 pairs of similar firms from the pool of 'potential ETS firms' and I flipped 272 coins to determine which firms, one in each pair, would become regulated. So 272 firms went from 'potential ETS firms' to being 'ETS firms,' and another 272 went from 'potential ETS firms' to 'non-ETS firms.' In January 2007 I created 6 additional pairs from the pool of 'potential ETS firms' and flipped 6 coins. In January 2008 I created 167 additional pairs and flipped 167 additional coins. The result is a set of 445 'ETS firms' and as many ex ante similar 'non-ETS firms.'

Next, imagine that someone accidentally deleted the list of paired controls in this experiment. So now I have to go back to the year each ETS firm became an 'ETS firm'—2005 or 2007 or 2008—and search through the set of 'potential ETS firms' for the one that was part of the same pair, and I call this a 'non-ETS firm.' Importantly, I exclude other ETS firms when I do this search, even the ones that weren't in the EU ETS yet. Since I know that they ended up as 'ETS firms,' I know they couldn't have been selected as 'non-ETS firms' in the experiment. This means that firms cannot transition from being controls to treated.

I didn't run this experiment, of course, and there was no list of controls to misplace. Matching does not reconstruct a true experimental data set, then, but rather creates a data set that looks as if it was generated by this experiment. The data set is made to look just like if treatment had been assigned by coin flips, even though it wasn't. Because the coin flips didn't actually take place, we have to worry that pairs of *ex ante* similar firms might be systematically different in some unobserved but important way. Fortunately, the EU ETS's design tells us that the treatment received by suitably matched companies is plausibly uncorrelated with their potential outcomes, just like in a true experiment.

The reason for this is that the EU ETS only regulates plants above certain activity-specific capacity thresholds. The pivotal difference between an ETS company and a non-ETS company, then, is whether or not its largest plant exceeds the threshold. Although I do not observe plant capacities,[2] I know that ETS companies must have at least one plant with capacity in excess of the threshold, and non-ETS companies must not. The research design in this paper is based on the assumption that, conditional on a set of key company-level characteristics including past adoption and innovation activity, the capacity of the largest plant a firm operates affects future adoption and innovation only through its role in determining ETS-status.[3] It follows from this that I can omit the plant capacity and compare *otherwise* identical ETS and non-ETS companies. Conditional on these other covariates, one cannot predict which firm will do more or less adoption and innovation in the future.

[2]The production capacities of plants are unobserved because records of plant capacities appear not to have been preserved. I have had several long conversations with British and European regulators to confirm this fact beyond a reasonable doubt.

[3]This assumption is untestable with the available data. I have searched for evidence both for and against this exclusion restriction in the published literature and found nothing. The absence of any significant theoretical or empirical work on the role of plant capacity in determining adoption and innovation does suggest that it probably isn't thought to be important, like the vast majority of unstudied predictors.

There is unfortunately no definitive way to determine which set of covariates one should use to judge whether companies have similar potential adoption and innovation outcomes. Based on the existing empirical literature, we would expect propensities for adoption and innovation to be quite different across economic sectors, across newer and older companies, and across larger and smaller companies, so it seems natural to want to condition on these basic company characteristics. Additionally, adoption and innovation are dynamic processes that exhibit a high degree of path-dependence, so past values of the outcome variables are probably informative about future outcomes. Future outcomes may also be a response to past non-ETS regulations, so it's important to condition on this information as well, if at all possible. Finally, since administrative databases tend to lack data more often for smaller, newer, less active companies (precisely the sort of companies believed to have different adoption and innovation propensities), it's important to condition on missingness rather than to treat it as a random occurrence. The hope is that, once we condition on all these covariates, future outcomes are expected to be as good as random in the absence of intervention.[4]

The matching design must also take account of variation in the timing of treatment. One possibility is that firms could manipulate the timing of entry into the EU ETS, which could undermine the assumption of conditional unconfoundedness. In particular, we might imagine that firms that find adoption and innovation especially difficult rushing to sign up for the UK ETS and CCAs in the hopes of forestalling entry into the more stringent EU ETS. Even if control firms signed up at the same time, they might be doing so because they expected to adoption and innovation to be especially easy, and that they could earn some rents by entering these programs. This would bias my estimates against finding an effect on directed technological change.

There are two important problems with this story. First, the main draw of CCAs was that signatories earned an 80% discount on their Climate Change Levy bill. This provided a strong incentive to sign up for CCAs whether or not a company was anticipating future regulations under the EU ETS, and indeed, the main driver of participation was eligibility (Martin et al., 2014). If we're worried that EU ETS firms signed up to avoid stricter climate regulations, it seems non-ETS firms signed up for exactly the same reason. Second, the

[4]Matching on these covariates should also tend to mitigate any lingering concerns about my exclusion restriction. A control company in the same sector, of similar age and size, with similar history is probably much more similar to the ETS company in terms of the unobserved plant capacity than a randomly chosen non-ETS company.

timing doesn't fit. The list of Direct Participants in the UK ETS was determined through an auction held in early 2002, but the related EU ETS exemption wasn't announced until 2003. The CCA exemption was also announced in 2003, but after that point there were no firms added to the CCA exclusion lists.[5] Given the incentives and timing, then, EU ETS firms wouldn't and couldn't have differentially self-selected into the UK ETS or CCAs in order to delay entry into the EU ETS.

There are three criteria we should use to evaluate the success of any particular matched sample. The first is the credibility of the claim that differences in treatment of *ex ante* similar firms are not driven by unobserved factors that are correlated with potential outcomes. I have tried to address this question already. Second, we want as high a degree of covariate balance as is possible. If the covariates are poorly balanced, it becomes impossible to tell whether the difference in outcomes measures the treatment's effect or the effect of some other influence. It is, at least potentially, a biased estimator. Third, we want to match as many ETS companies as possible. Covariate balance can be trivially maximised by matching a single pair, but statistical power clearly diminishes the fewer matched pairs we have. The fewer pairs, the higher the variance of the resulting estimates. So we care about having a *large* and *balanced* matched sample. Greater balance eliminates potential sources of bias, while greater sample size keeps the variance in check. We ultimately have to make a judgement call on how to balance these conflicting objectives, and it will be prudent to check that the results are not too sensitive to the exact choices we make.

Computationally, matching was implemented in three steps. In the first step, I used the coarsened exact matching algorithm developed by Iacus et al. (2012) to retain only those non-ETS companies that are close enough to at least one ETS company to have any chance of being matched to it. Given the large number of firms and covariates I start off with, it was computationally necessary to quickly and efficiently discard the overwhelming number of poor candidates for matching before more sensitive (and computationally demanding) balancing techniques could be applied.

In the second step, I used the R-function `GenMatch` to construct a matched set (Sekhon, 2007). It uses a genetic search algorithm to automatically canvas the space of generalised Mahalanobis distance metrics until it finds one that maximises covariate balance. This process automates the choice of covariate weights, taking it out of the researcher's hands. This method also has the advantage of producing more reliable and balanced outputs

[5]Based on personal communication with Environment Agency staff responsible for implementing the EU ETS at the time.

than regression-based propensity score models when the treated group is a particularly unrepresentative sample of the population, when the covariate distributions are highly skewed, span several orders of magnitude, and contain many missing values. I executed this search without imposing any calipers beyond the exact-match restriction on economic sector. This ensures that every ETS company will be matched to at least one non-ETS company, whereas imposing stricter calipers up front would have resulted in automatically discarding some number of ETS companies without locating the best match for them. This yields a matched sample of 445 ETS companies and 473 non-ETS companies. I take the average whenever I find more than one match so that there is a weighted average control firm matched to each treated one.

In the third and final step, I iteratively excluded pairs that most impair covariate balance until reasonable balance is achieved. This resulted in the removal of 42 ETS companies for which the best matches were relatively poor. This third step amounts to imposing calipers after-the-match, and though this may entail some loss in efficiency, it guarantees that the final matched sample is a strict subset of the full matched sample. This makes it much easier to examine sensitivity with respect to how sample size and balance are traded off. Rather than re-running the whole expensive matching algorithm for an arbitrary number of caliper choices, it is enough to check how sensitive the conclusions are to the omission or inclusion those 42 pairs. Omitting them amounts to putting a premium on covariate balance, while including them amounts to desiring the largest possible sample.

An important advantage of this matching algorithm is that it provides some assurance that there isn't a larger more balanced sample buried among the mass of discarded observations. Whether the sample I've got is large and balanced enough is a judgement call, but the matching algorithm at least assures us that it's the best there is.

The design and execution of this matching algorithm have been informed by previous research on the determinants of adoption and innovation, as well as by computational requirements and constraints. But aside from these *ex ante* justifications, at this stage an important reason to prefer my particular matching method and constellation of covariates is that I selected them (and the resulting matches) prior to viewing any of the post-ETS outcomes. The selection of matches therefore could not have been influenced (even unintentionally) by the results they would turn out to produce. Below I will discuss the results from a number of alternative specifications suggested by colleagues, conference participants, and anonymous referees. But since both I and they had seen the main results before devising these robustness checks, I cannot offer the same iron-clad guarantee as I am prepared to offer for the main results.

IV. Additional robustness checks

A. *Sensitivity to outliers*

Carbon intensity, patenting, and R&D spending all have highly skewed distributions, and this raises the possibility that my estimates might be driven by one or a few highly leveraged pairs. To examine the sensitivity of my estimates, table 2 reports the re-estimated treatment effects obtained after excluding the farthest outliers for each outcome variable. Note that an outlier can lie at either the top or bottom of the distribution, so excluding it can either lower or raise the estimate. Relatedly, the reduction in variance that results from its exclusion can either outweigh the reduction in sample size or not, so the statistical significance can increase or decrease when outliers are omitted.

Table 2: Sensitivity to outliers

	Main estimate	Leave one out	Leave five out
Efficiency			
CO_2 intensity (tCO_2 / £1,000)	0.078	0.065	0.025
	(0.063)	(0.135)	(0.268)
Labor intensity (employees / £1,000)	-0.002	-0.002	0.002
	(<0.001)	(<0.001)	(<0.001)
Patenting			
Low-carbon patents	0.415	0.475	0.325
	(0.079)	(0.034)	(0.063)
All other patents	0.130	0.130	0.130
	(0.167)	(0.203)	(0.190)
R&D spending (£1,000s)			
Low-carbon R&D	200.000	200.000	65.000
	(0.046)	(0.046)	(0.084)
Total R&D	514.000	496.000	496.000
	(<0.001)	(<0.001)	(<0.001)

Notes: The parenthetical number below each estimate is the p-value associated with a test of the null hypothesis of zero effect.

Most notably, the estimate on CO_2 intensity falls substantially in both magnitude and statistical significance as outliers are excluded. The estimate on low-carbon patents is far less sensitive, with the point estimates corresponding to a range of 25 to 30 additional low-carbon patents. Moreover, excluding outliers increases statistical significance of the point estimates on low-carbon patents. Data on low-carbon R&D is missing for most firms, and excluding more than just one or two outliers results in a substantial drop in the point estimate, though the statistical significance is only moderately impaired. The estimates on labour intensity, non-low-carbon patents, and total R&D spending are all insensitive to outliers.

These re-estimated treatment effects provide a reason for being cautious about inferring that the EU ETS has caused an increase in CO_2 intensity. They also emphasize the limitations of the available data on low-carbon R&D. We can have reasonable confidence that the effect is indeed positive, but need to be more cautious in interpreting the economic magnitude of this effect.

B. *Full sample estimates*

To obtain a sufficiently balanced matched sample, I ended up excluding 42 ETS firms for which I could only find relatively poor matches. I might insist, at the other extreme, on keeping all of the ETS firms no matter the cost to balance. This exposes me to greater risk of conflating the EU ETS's effects with other factors, but instead eliminates the potential bias introduced by selection. As a matter of completeness, I re-estimate the treatment effects here without omitting these 42 pairs.

Table 3 and figure 2 show the pre-ETS covariate balance for the full matched sample of 445 EU ETS companies and 473 non-EU ETS companies. The distribution of covariates is unbalanced in several important dimensions. The regulated companies are more than twice the size of their unregulated counterparts. The differences are not statistically significant only because of the inflated standard deviation associated with adding in the outliers. Under the equivalence ranges used for main analysis in the paper, these differences are highly statistically significant. Even with the more forgiving standard used here, there is strong statistical evidence that the two groups of companies have different pre-treatment CO_2-intensities, patenting, R&D spending, and regulatory histories. These differences are obviously driven by the 42 pairs that were omitted from the main analysis.

While potential outcomes are plausibly independent of treatment for the sub-sample used in the main analysis, the full sample appears to include regulated companies that would not plausibly have been left unregulated in any incarnation of the EU ETS. This makes it difficult to think of a reasonable basis for constructing empirical counterfactuals for these companies, which is why they were omitted from the main analysis.

The limitations associated with analysing the full sample are now obvious, but even

Table 3: Equivalence tests for all EU ETS ($N = 445$) and matched non-EU ETS companies ($N = 473$)

	EU ETS mean	Non-EU ETS mean	Equivalence range	Signed-rank test p-value	Paired t-test p-value
Company basics					
Revenues (£1,000s)	754,417	420,527	±543,983	<0.001	<0.001
Employees	2,281	1,038	±1,216	<0.001	0.473–
Year of birth	1987	1988	±2	<0.001	<0.001
Economic sector (3-digit)	*Exact match*	–	–	–	–
Efficiency					
CO_2 intensity	0.398	0.232	±0.103	0.531	0.892
Labour intensity	0.011	0.012	±0.002	0.002	0.003
Patenting					
Low-carbon patents	0.113	0.014	±0.160	0.994^{LP}	0.127
All other patents	2.603	0.723	±2.284	0.010	0.243
R&D spending					
Low-carbon patent share	0.010	0.002	±0.009	0.999^{LP}	0.423
R&D spending (£1,000s)	19,758	10,795	±13,267	<0.001	0.126
Regulatory history					
CCA participation	0.508	0.387	±0.049	–	–
UK ETS participation	*Undisclosed*	*Undisclosed*	–	–	–
CCL bill (£1,000)	107	67	±33	0.620	0.622
R&D support (£1,000)	1,158	1,411	±2,047	<0.001	<0.001

so, it may be reassuring to see that the substance of my earlier conclusions are not overly sensitive to this modification. Figure 3 and table 4 show the outcomes for the full sample of matched companies. The trends and estimates are qualitatively consistent with those reported in the main analysis, although the magnitudes are generally larger and the p-values smaller. As I have outlined earlier, however, the case for attributing these larger estimates to the EU ETS is much more problematic. They are reported here primarily for completeness.

C. *Alternative matching*

The algorithm I have used to obtain matches already anticipates a number issues. First, the `GenMatch`-function executes an automated search of the space of covariate weights, ensuring that choosing differently will produce inferior balance. It isn't clear what, if anything, can be learned from choosing a different and arbitrary set of weights. Second, because I didn't impose strict calipers up front, I have been able to re-estimate my results

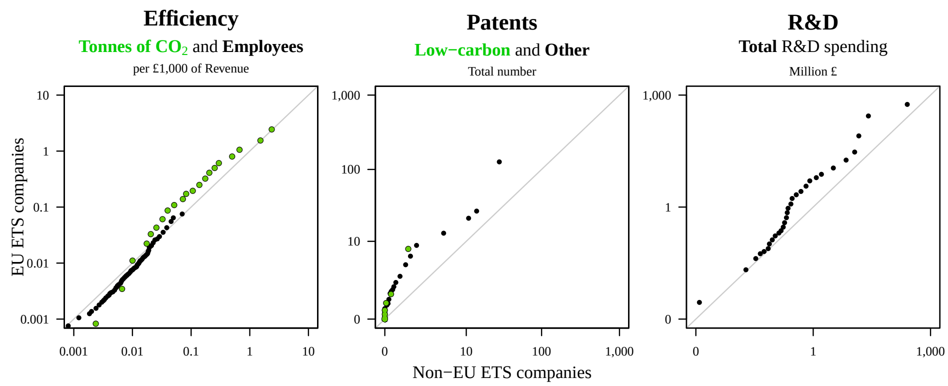

Figure 2: Comparison of full sample of matched EU ETS and non-EU ETS companies

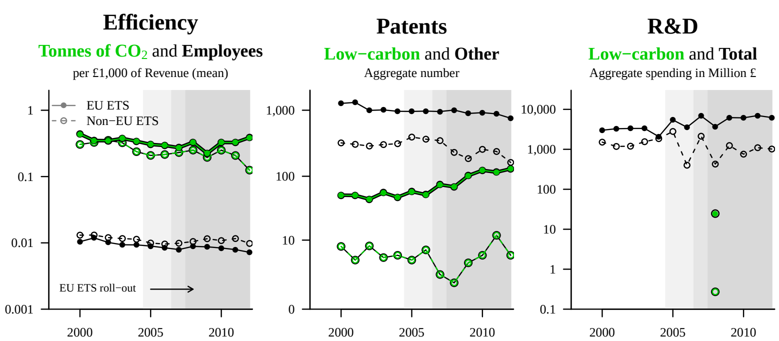

Figure 3: Adoption and innovation of all matched EU ETS and non-EU ETS companies

Table 4: Matching estimates of the effects of the EU ETS

	Hodges-Lehmann point estimate	*p*-value
Efficiency		
CO_2-intensity	0.089	0.008
Labour intensity	-0.002	<0.001
Patenting		
Low-carbon patents	0.680	<0.001
Other patents	0.740	<0.001
R&D spending (£1,000s)		
Low-carbon R&D	521.000	0.181
Total R&D	1,024.000	<0.001

for the full matched set, of which the main sample is a strict subset. This exercise is equivalent to loosening all of the implicit calipers until none are binding. All except one, that is. The main researcher-margin left to examine is my choice to match exactly on economic sector as defined at the 3-digit level rather than at the more disaggregated 4-digit level. Requiring matched pairs to be drawn from within the same 4-digit sector naturally reduces the number of potential matches, and will in expectation result in a deterioration of balance on other covariates.

Table 5 reports the pre-ETS covariate balance for the same 403 ETS companies used in the main analysis, but this time matched to 462 non-ETS companies drawn from within the same economic sectors as defined at the 4-digit level.[6] On measures of efficiency, too, this matched sample is more balanced than the sample used in the main analysis. But as expected, there are more substantial imbalances on a range of other covariates. The disparities in size, patenting, R&D spending, and past policies, are all greater.

Table 6 reports the estimates for this matched sample.[7] The biggest change is in the estimated effect on CO_2-intensity. This is also the outcome we should pay the most attention to here, because this is the main covariate for which the current sample achieves greater balance than the sample used in the main analysis. The revised estimate is a great deal smaller than the original, and statistically indistinguishable from zero. The confidence interval admits a somewhat greater chance of adoption than before, but the most likely outcome is still that there hasn't been significant adoption. This alternative estimate hardly contradicts the interpretation that my results should be read primarily as evidence against widespread adoption.

The changes to the estimated effects on patenting and R&D spending are perhaps less informative, given the greater pre-treatment imbalances in these variables (and the greater imbalance in R&D support). The new estimates suggest a somewhat smaller effect on low-carbon patenting but a larger effect on R&D spending. Neither estimate is significantly different from the original estimate, nor different from zero. The signs of the point estimates haven't changed, and they still display the same general pattern of companies emphasising low-carbon innovation.

It might have been alarming if re-matching within the same 4-digit economic sectors

[6] The quantile-quantile plots for the outcome variables are available in the replication archive at the UK Data Service, but have not been approved for public release for reasons pertaining to rules of disclosure.

[7] An accompanying time-series plot for the outcome variables is available in the replication archive at the UK Data Service, but has not been approved for public release for reasons pertaining to rules of disclosure.

Table 5: Equivalence tests for EU ETS ($N = 403$) and non-EU ETS companies ($N = 462$) matched at the 4-digit level

	EU ETS mean	Non-EU ETS mean	Equivalence range	Signed-rank test p-value	Paired t-test p-value
Company basics					
Revenues (£1,000s)	520,473	343,728	±408,320	<0.001	<0.001
Employees	1,595	900	±766	<0.001	0.329
Year of birth	1988	1988	±2	<0.001	0.002
Economic sector (4-digit)	*Exact match*	–	–	–	–
Efficiency					
CO_2 intensity	0.292	0.294	±0.099	0.021	0.020
Labour intensity	0.011	0.011	±0.002	<0.001	<0.001
Patenting					
Low-carbon patents	0.019	0.005	±0.020	0.997^{LP}	0.123
All other patents	0.613	0.273	±0.440	0.738	0.163
R&D spending					
Low-carbon patent share	0.004	0.002	±0.007	0.990^{LP}	0.017
R&D spending (£1,000s)	10,193	2,354	±6,264	<0.001	0.447
Regulatory history					
CCA participation	0.479	0.340	±0.098	–	0.982
UK ETS participation	*Undisclosed*	*Undisclosed*	–	–	–
CCL bill (£1,000)	98	67	±27	0.210	0.539
R&D support (£1,000)	845	228	±1,060	<0.001	<0.3333

Table 6: Matching estimates of the effects of the EU ETS based on sample matched at the 4-digit level

	Hodges-Lehmann point estimate	p-value
Efficiency		
CO_2-intensity	0.029	0.191
Labour intensity	-0.002	<0.001
Patenting		
Low-carbon patents	0.155	0.198
Other patents	0.630	<0.001
R&D spending (£1,000s)		
Low-carbon R&D	334.500	0.158
Total R&D	2,144.000	<0.001

gave rise to a radically different set of estimates. As it is, the estimates are different enough to raise questions about the precise magnitude of the EU ETS's effect. But they are also qualitatively similar to my original estimates and do not suggest a need to revise the central conclusions of the paper.

D. *Alternative estimators*

The Hodges-Lehmann estimator is a standard estimator in many disciplines (e.g. medicine and epidemiology), and recommends itself here because its statistical logic maps directly onto the identifying assumption in a matched study design, for its ability to cope with data-censoring without needing restrictive or implausible assumptions on first-stage regressions, and for its superior statistical power and robustness with highly non-Normal outcome distributions. Even so, it may be reassuring to know that the general pattern of findings is broadly stable with some more familiar estimators.

Table 7 reports two of the most familiar estimators: the difference-in-means and the differences-in-differences. The estimated effects on CO_2-intensity and labour-intensity, where censoring is not a concern, are quantitatively close to the Hodges-Lehmann estimates. The p-values are higher since these estimators are less powerful. The estimates on patenting are closer to zero than the original estimates, as would be expected from the lack of adjustment for censoring. But when converted into comparable units, the new estimates are actually larger than the original ones. If I multiply out a 0.035 annual average increase across the 3,047 total treated firm-years, I end up with roughly 107 additional low-carbon patents, compared with my original estimate of 64 additional low-carbon patents. Similarly, the new estimates imply roughly 420 additional non-low-carbon patents, compared with my original estimate of 90. The new estimates suggest a larger absolute effect on low-carbon innovation, though the emphasis on low-carbon technologies is somewhat weaker than before. The same can be said of the estimated effects on low-carbon R&D.

Since I am only looking at simple difference-in-means here, it is straightforward to apply regression adjustment. In particular, the matched sample is slightly unbalanced with respect to participation in pre-ETS regulations. If these imbalances are correlated with the outcomes, I might have misattributed some part of their effect to the EU ETS. Table 8 therefore reports estimates based on a regression where outcomes are modelled as a function of EU ETS status as well as the pre-ETS policy variables. There is relatively little change in the magnitude of the estimates compared with table 7, and of course, the estimates are statistically indistinguishable at conventional significance levels. The general pattern of effects is stable, so it seems that the differences in past regulatory treatment

Table 7: Mean differences and difference-in-differences

	Difference-in-means	Difference-in-differences
Efficiency		
CO_2 intensity	0.111	0.076
	(0.031)	(0.150)
Labour intensity	-0.002	0.000
	(0.009)	(0.994)
Patenting		
Low-carbon patents	0.038	0.035
	(0.251)	(0.246)
All other patents	0.135	0.138
	(0.429)	(0.392)
R&D spending (£1,000s)		
Low-carbon R&D	58.000	–
	(0.183)	
Total R&D	3,269.000	4,791.000
	(0.229)	(0.303)

Notes: The parenthetical number below each estimate is the associated p-value from paired t-test with a null hypothesis of zero difference. It has not been possible to estimate the difference-in-differences for low-carbon R&D since this variable is only observed during the post-treatment period.

do not explain variation in outcomes over-and-above the different regulatory treatment received under the EU ETS.

Table 8: Regression estimates adjusting for pre-EU ETS differences in regulation

	Regression point estimate	p-value
Efficiency		
CO_2 intensity	0.154	0.025
Labour intensity	-0.002	0.049
Patenting		
Low-carbon patents	0.035	0.291
All other patents	0.104	0.599
R&D spending (£1,000s)		
Low-carbon R&D	137.000	0.188
Total R&D	1,890.000	0.684

References

Iacus, S., King, G., and Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political analysis*, 20(1):1–24.

Martin, R., De Preux, L. B., and Wagner, U. J. (2014). The impact of a carbon tax on manufacturing: Evidence from microdata. *Journal of Public Economics*, 117:1–14.

Sekhon, J. (2007). Multivariate and propensity score matching software with automated balance optimization: The matching package for R. *Journal of Statistical Software*, 10(2):1–51.