

Bremer, Leon; Sommer, Konstantin

**Working Paper**

## Competitiveness and investments under emissions trading

Tinbergen Institute Discussion Paper, No. TI 2022-061/V

**Provided in Cooperation with:**

Tinbergen Institute, Amsterdam and Rotterdam

*Suggested Citation:* Bremer, Leon; Sommer, Konstantin (2022) : Competitiveness and investments under emissions trading, Tinbergen Institute Discussion Paper, No. TI 2022-061/V, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:

<https://hdl.handle.net/10419/265837>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

TI 2022-061/V  
Tinbergen Institute Discussion Paper

# Competitiveness and investments under emissions trading

*Leon Bremer<sup>1</sup>*  
*Konstantin Sommer<sup>1</sup>*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and Vrije Universiteit Amsterdam.

Contact: [discussionpapers@tinbergen.nl](mailto:discussionpapers@tinbergen.nl)

More TI discussion papers can be downloaded at <https://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam  
Gustav Mahlerplein 117  
1082 MS Amsterdam  
The Netherlands  
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam  
Burg. Oudlaan 50  
3062 PA Rotterdam  
The Netherlands  
Tel.: +31(0)10 408 8900

# Competitiveness and investments under emissions trading

Leon Bremer<sup>\*†</sup>      Konstantin Sommer<sup>†‡</sup>

September 2, 2022

## Abstract

We study the effects of the EU Emissions Trading System (ETS) on employment and profits as well as on the investment decisions of Dutch manufacturing firms. Motivated both by sizable differences between firms that are regulated in different phases and by the gradual increase in regulatory stringency, we pay close attention to treatment effect heterogeneity between firms and over time. We use microdata from Statistics Netherlands to apply two difference-in-differences (DiD) estimators: (1) a matched two-way fixed effects regression and (2) a recently developed, more flexible DiD method, designed for staggered treatment and treatment effect heterogeneity. We find that firms that were first regulated in phase 1 and 2 experience temporary employment losses of between 7 to 9% early in the regulation, but we do not find conclusive evidence for changes in profits. Firms that were regulated the earliest reduced their investments throughout all phases.

**JEL codes:** H23, L51, Q52

**Keywords:** Emissions trading, Environmental regulation, Staggered Difference-in-Differences, Treatment heterogeneity, Manufacturing

---

\*Vrije Universiteit Amsterdam, Department of Spatial Economics, De Boelelaan 1105, 1081 HV Amsterdam, and Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam. E-mail: [l.bremer@vu.nl](mailto:l.bremer@vu.nl).

†University of Amsterdam, Department of Macro and International Economics, Roetersstraat 11, 1018 WB Amsterdam, and Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam. E-mail: [k.h.l.sommer2@uva.nl](mailto:k.h.l.sommer2@uva.nl). (corresponding author)

‡The authors are grateful for funding by the A Sustainable Future (ASF) platform from the University of Amsterdam and the travel grant provided by the International Association for Applied Econometrics (IAAE) to its 2022 conference. The authors are also grateful for the opportunities to present at various conferences and seminars, and the feedback received by various participants of these sessions.



# 1 Introduction

This paper studies the effects of the European Union Emissions Trading System (EU ETS) on both the competitiveness as well as the investment decisions of regulated manufacturing firms. We show that firms that became regulated in different years differ from each other, and that regulation stringency significantly varied over time. We carefully disentangle the effects on these different cohorts of firms, and analyse the dynamics of the policy’s effects over time.

The EU ETS is the world’s largest cap-and-trade system, aiming to reduce the EU’s emissions of greenhouse gases. It was implemented in 2005 and caps the amount of emissions within the EU for the covered plants, which together make up for about half of the EU’s emissions. Large emitters have to surrender one allowance for each emitted ton of CO<sub>2</sub>-equivalent at the end of each year. These allowances can be traded on financial markets, thereby establishing a price for carbon. ETS regulation has been amended throughout its four phases (2005-07, 2008-12, 2013-20, 2021-30), in each of which additional installations were regulated and in which the regulation and its stringency were adapted.

Since the beginning of the ETS, policy makers and industry representatives have expressed concerns about the problem of (carbon) leakage. This would occur if EU producers lose competitiveness towards their non-EU competitors due to compliance with the ETS, leading to a loss of jobs in Europe while not reducing global emissions, but merely shifting them elsewhere. This would be undesirable both from an EU welfare, and from an environmental point of view. Meanwhile, stricter regulation could also incentivize firms to invest in new technologies and gain competitiveness in the long run, a hypothesis in line with Porter and Van der Linde (1995). This study will estimate the causal effect of the ETS on both of these potential policy side effects.

While previous studies have estimated the effects of the ETS, few analyzed more than one dimension of the underlying treatment effect heterogeneity. We focus our analysis on these heterogeneities, motivated by the differences between firms that start being regulated in different phases (which we will call “cohorts”) and by the changing nature of the regulation over time. If the ETS’s effect differs between regulated cohorts, neglecting such heterogeneities might lead to biased estimates that could explain the absence of significant findings in the literature.

Furthermore, few studies cover the more recent years in which allowance prices rose and in which amendments like the Market Stability Reserve (MSR) were introduced and implemented. This paper tries to close these gaps.

We are able to use detailed firm-level microdata from Statistics Netherlands (in Dutch: CBS), the Dutch national statistics agency, and link those to the European Union Transaction Log (EUTL) for information on regulated ETS firms. Besides the more classic difference-in-differences (DiD) implementation through matched

two-way fixed effects (TWFE), we employ a recent, more flexible DiD method developed in Callaway and Sant’Anna (2021) (which we will refer to as “CS”) that is designed for settings with staggered treatment timing, like ours. The latter method is motivated by a recent stream of econometric literature revealing several flaws of TWFE estimators in staggered settings with treatment heterogeneity.

In both approaches, we aim to disentangle the average treatment effect on the treated (ATT) for each cohort-phase combination. This offers ways to break down the effects of the EU ETS between its phases, as well as between its participants. As cohorts could differ from each other, it could imply that these firms also exhibit different treatment effects from the regulation. It also seems likely that effects are not constant over time, as firms might adjust to the regulation and because the nature of the ETS has changed over time, especially in between phases.

The preferred CS method shows negative immediate, but temporary, effects from the EU ETS on employment. Both cohort 1 and cohort 2 experience negative employment effects of between 6-7% in their first regulated phase, but they return to normal employment levels afterward.

Investments also seem to suffer under ETS regulation, as cohort 1 firms persistently reduce their investments with 0.32 to 0.47 industry-level standard deviations. For profitability, the results are less conclusive and mostly statistically insignificant. It thus seems that the large firms belonging to cohort 1 and 2 reduce their size without compromising their profit margins.

This implies that, against our intuition, the last phase, which was the most stringent so far, did not lead to larger losses in competitiveness. The fact that we nevertheless establish these negative effects on firms’ investments and (temporarily) on employment could be driven by the export and energy intensity of the Dutch manufacturing sector.

The two methods result in somewhat different findings, as the TWFE regression indicates larger employment losses for the first cohort, but does not indicate significant reductions in investments. These differences are likely to be driven by the fact that the control group in the CS estimation is a better fit, and highlight the importance of the right methodology choice in staggered DiD settings. Testing for the robustness of our results and discussing the underlying assumptions confirms our results.

The paper continues as follows: [Section 2](#) discusses the related literature. The data and policy background are discussed in [section 3](#). The methodology and results are presented in [section 4](#) and [section 5](#) respectively. [Section 6](#) shows the robustness of these results and discusses the assumptions underlying our identification. [Section 7](#) concludes this research.

## 2 Related literature

There are several studies on the effects of the ETS on the competitiveness of regulated firms as well as potentially resulting carbon leakage and related technology adoption. Several of these studies use administrative firm-level data in other countries and apply comparable difference-in-differences methods. Other studies utilize a larger set of EU ETS firms combined with publicly available data sets (e.g. Calel & Dechezleprêtre, 2016). Underlying all studies is the complexity of finding appropriate control firms that are unregulated but sufficiently similar, in order to draw causal conclusions.

Most studies that rely on matching estimate the treatment effect on the treated by using the semi-parametric estimator of Heckman et al. (1997). In this framework Petrick and Wagner (2014) and Jaraite-Kažukauske and Di Maria (2016) find no negative effects of the ETS on productivity and employment for Germany and Lithuania, respectively, and Colmer et al. (2022) find no evidence for outsourcing in France. Löschel et al. (2019) additionally use a two-way fixed effects setting to analyze the ETS's effect on productivity in Germany. The authors interestingly find significant positive effects on productivity using the Heckman-style estimator, but not in the regression estimation. According to the authors this effect is likely driven by a positive EU ETS effect on efficiency in some of the regulated industries. Marin et al. (2018), using non-administrative micro data from Bureau van Dijk for a larger set of countries, also do not find negative effects on economic performance, but do find an increase in labor productivity.

All of these studies, however, only use data for the first phase and some years into the second phase. Since the stringency of the ETS increased significantly in the second and third phases, adding later phases may lead to stronger and more clear-cut results. Two studies that do look at later phases, however, also do not find strong negative effects of ETS regulation. Dechezleprêtre et al. (2019) use data on multinational firms up to 2014 and analyze carbon shifting within these firms, again without finding much evidence of leakage. Klemetsen et al. (2020) do look at phase 3 and analyze firms in Norway. Their regression methodology accounts for different effects between phases, but not between companies starting in different phases. They find a slight increase in productivity in phase 2, but no significant effect in the other phases.

In a literature survey, Verde (2020) comes to the conclusion that there is no convincing evidence of leakage and losses in competitiveness due to the ETS yet. The authors also highlight that this might be due to the short time span covered in almost all studies and point to the importance of analyzing more long-term indicators like investments. This study is trying to address both of these gaps.

When it comes to the ETS's effect on innovation, the literature is smaller, but still contains important contributions. Calel and Dechezleprêtre (2016) show

in a large multi-country panel that the ETS has increased green patenting, and Borghesi et al. (2015) show in an Italian phase 1 firm-level panel that regulated sectors have increased innovation, but that this varied by treatment stringency of the sector. A survey by Teixidó et al. (2019), however, comes to the conclusion that evidence on the ETS’s effect on innovation is still too sparse for a coherent conclusion.

Further, to address time-varying heterogeneity in the treatment effects the econometric design needs to be appropriate. There is recent discussion of DiD and its potentially imperfect approximation using TWFE (see e.g. de Chaisemartin & D’Haultfoeuille, 2022). Callaway and Sant’Anna (2021) offer a more flexible alternative to TWFE. In this paper we will employ their estimator on top of the more classical TWFE estimator.

Our contribution to the previous work is threefold. First, we add analytically to the debate about causal effects of the ETS, by adding insights into the potential treatment heterogeneity of the ETS. For this, we employ both a classical matched-TWFE estimator, but adapt it such that it accommodates cohort-phase specific estimates, as well as a recently-developed more detailed semi-parametric estimator. Second, we benefit from longer time series, allowing us to estimate the later phase’s effects, as well as the longer term effects from earlier phases. Third, we are able to use detailed administrative data. We have access to data on investments, on top of the more traditional indicators for firm performance and competitiveness. Additionally, the Netherlands are due to their export orientation and rather energy-intensive industrial structure a country in which competitiveness effects might be more clearly visible.

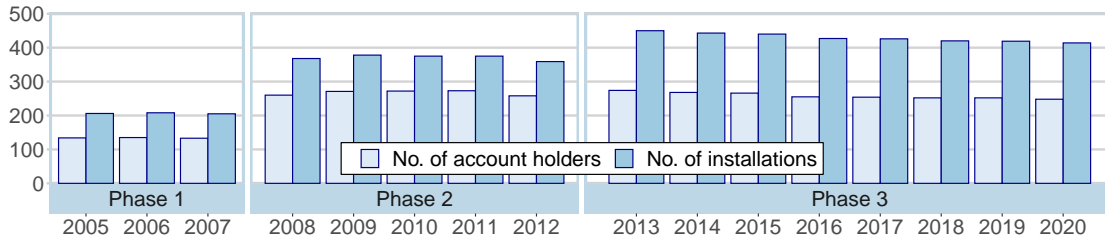
## 3 Data and policy background

### 3.1 EU ETS policy background

The EU ETS regulates installations, which we will also refer to as plants. Each of these plants is registered under one owner, the account holder, at a time in the European Union Transaction Log. The amount of regulated active installations in the Netherlands and their account holders can be found in [Figure 1](#).<sup>1</sup> After its initial implementation in 2005 the ETS has been largely revised 3 times when new phases came into effect, in 2008, 2013 and 2021. Most of these revisions aimed at making the system more restrictive and effective. This study uses data until 2020, thereby excluding Phase 4.

---

<sup>1</sup>Only (former) Operator Holding Accounts that are registered in the Netherlands are selected. The connected installation must have positive verified emissions for that year.



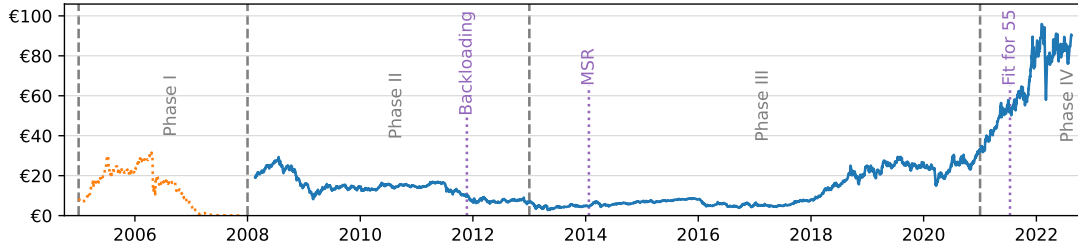
**Figure 1:** *Account holders and installations regulated under the EU ETS.*

*Note:* Number of active installations regulated under the EU ETS in the Netherlands and their account holders. Source: authors calculations based on EUTL data accessed through [EUETS.INFO](http://EUETS.INFO).

In phase 1 (2005-2007) allowances were handed out so plentiful that their price dropped to zero towards the end of the phase, see [Figure 2](#). In the Netherlands actual emissions were almost 15 percent below the number of allocated allowances (Ellerman & Buchner, 2008). Phase 2 (2008-2012) added nitrous oxide as a greenhouse gas and increased the penalty for non-compliance from €40 to €100 per tonne of CO<sub>2</sub>-equivalent. The amount of regulated installations within the Netherlands increased from 205 to 368 (see [Figure 1](#)), mostly because in Phase 1 150 Dutch installations were excluded from the ETS.<sup>2</sup>

Even more greenhouse gases were added in Phase 3 (2013-2020). Also, the default allowance allocation method switched from grandfathering to auctioning. To counteract the low emissions prices that were arguably caused by sustained low demand, the European Commission implemented two sets of new rules to the ETS. First, starting from 2014 the auction of new allowances was postponed until 2019-2020, which was referred to as Backloading. Second, in 2019 the Market Stability Reserve (MSR) started operating. The MSR takes the backloaded allowances and puts them in a reserve. Depending on demand and supply, allowances will be added to the reserve or released from the reserve. As of 2023 excess allowances in the reserve might be permanently cancelled. Further, manufacturing sectors in the aluminium and chemicals production were added to the coverage. This did not change the number of regulated account holders much, but it significantly increased the number of regulated plants (see [Figure 1](#)). Arguably, more plants of the same owners were regulated in phase 3.

<sup>2</sup>The following decisions by the European Commission (EC) provide further details of the phase 1 exemptions for Dutch installations. In October 2004 the EC exempted 93 installations and in March 2005 the EC exempted a further 57 installations (European Commission, 2004, 2005). Other countries that have exempted some firms from the regulation in the first phase were the UK, Sweden and Belgium. More information can be found on the [EU Commission's website](http://EUCommission's website)



**Figure 2:** *EU ETS allowance price.*

*Note: The EU ETS's allowance price in Euros per tonne of CO<sub>2</sub>-equivalent. These are day closing prices for its futures contracts. The futures montage ECF00-NDEX is plotted in solid blue, and this data is accessed through FactSet. The December 2007 futures price for phase 1 allowances is plotted as a dotted orange line. These allowances were not transferable to later phases. The phase 1 data come from the [European Environment Agency](#). Vertical dashed grey lines indicate the starts of a new phase, while purple dotted vertical lines indicate early proposal dates of amendments to the EU ETS.*

Phase 4 (2021-2030) mainly sped up the rate at which the cap decreases over time, the Linear Reduction Factor (LRF), and it strengthened the MSR.<sup>3</sup>

The changing degree of stringency is also reflected in the allowance price path, as depicted in [Figure 2](#).<sup>4</sup> Prices decreased to zero at the end of phase 1, then started around €20 in phase 2, but stayed around only €10 for several years. Even though economists argue about the optimal price of carbon, such low prices have almost uniformly been deemed as too low to have the intended impact. Prices have started to increase since 2017/2018 and have nearly reached €100 in 2022, making the ETS far more restrictive in recent years.

### 3.2 EUTL and Dutch microdata

The data for this project comes from two main sources. First, the European Union Transactions Log data is accessed through [EUETS.INFO](#), a free service that provides cleaned data from the EUTL (Abrell, 2021). Second, Dutch firm-level data is accessed through the microdata services of Statistics Netherlands (in Dutch: CBS).

The data collected from the EUTL contain information on the free allocations of allowances, verified emissions, allowances surrendered, and the use of interna-

<sup>3</sup>Please refer to the [European Commission's webpage](#) for more details.

<sup>4</sup>ETS stringency is not the only driver of the allowance price. A body of literature studying the ETS price drivers has identified fossil fuel prices to play a key role (see e.g. Hintermann, 2010).

tional credits, both by installation and account holder. Account holders in the EUTL can potentially own several regulated plants and are registered under a national identification number.<sup>5</sup> The data are organized in an unbalanced panel spanning the years 2005-2020 and a total of 439 unique account holders, owning 598 installations.

The CBS data are not publicly accessible and are anonymized. They contain rich firm-level information on economic activity of almost the entire population of Dutch firms with more than 50 employees. The data contain information like the number of employees, costs of goods sold and turnover, as well as investment data. This study is restricted in scope to manufacturing firms and relies on more than 40,000 firms over a time span of 21 years. To deflate monetary variables, we use Eurostat’s industry producer price index for the Netherlands.

We link EUTL data to the administrative firm-level data of CBS. The linking takes place by the use of the chamber of commerce identifiers that are available in the EUTL for Dutch account holders and in the CBS data. Within CBS, several chamber of commerce numbers can comprise a “business unit”, a construct defined by CBS and further explained in [Appendix A.3](#). We will from here on refer to these business units as “firms”. After linking the EUTL data to CBS’s anonymized data we are not able to identify individual firms anymore.

As a business unit can comprise multiple account holders and plants, it can be the case that a business unit is regulated through more than one plant. We do not make a distinction here and consider each business unit (firm) as regulated if it owns at least one regulated plant in that year. Our level of analysis is on this business unit level, referred to as the firm level.

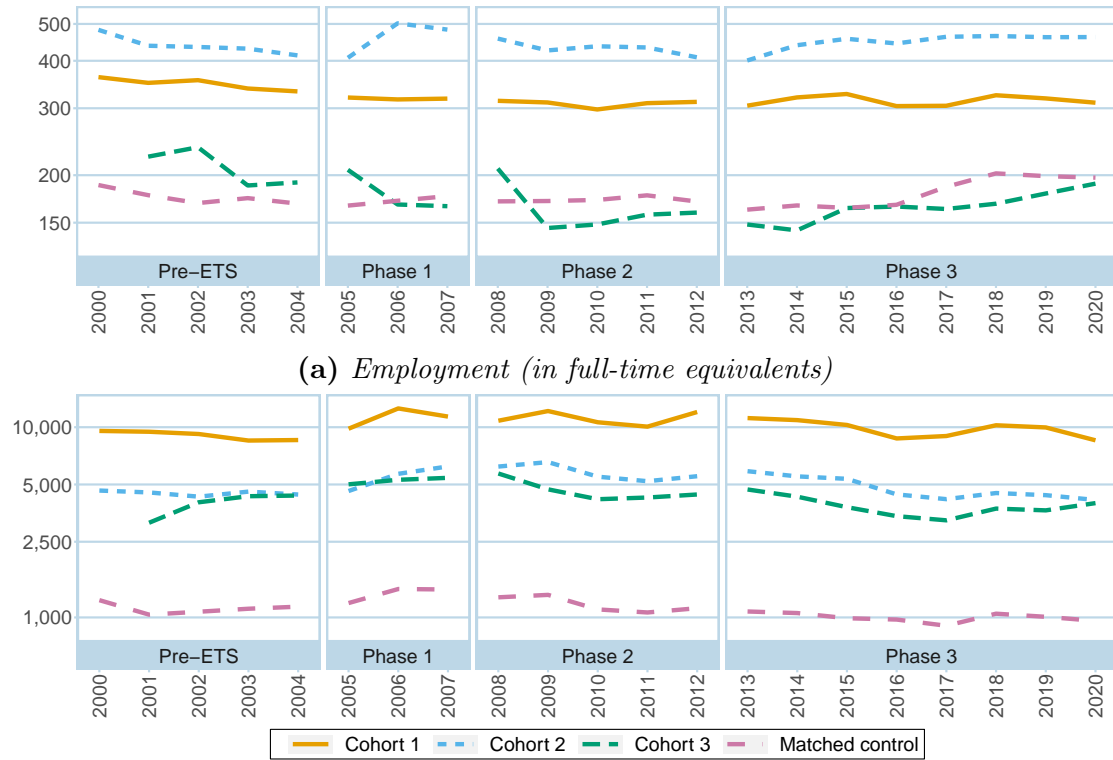
### 3.3 Descriptives

In this section we will elaborate on the two most important sources of heterogeneity that this study tries to disentangle. First, we show the substantial differences between the regulated cohorts, and then we show the development of the ETS treatment stringency over time.

[Figure 3](#) shows the development of the average firm over time for energy expenditure and employment. The plot shows averages for the different ETS cohorts as well as for a set of matched control firms, that is chosen to be as similar to the treated firms as possible, as is further outlined in [Appendix C](#).

---

<sup>5</sup>As installations are assets, they can be purchased from or transferred to other firms. Such changes of ownership are not perfectly captured by the data. Many installations do not change ownership between EU ETS phases in our data, but for the ones where it does change, we manually look up the date of ownership change using online public sources. Sources can be online news articles or websites that provide information about ownership structures. The list of manually assigned ownership changes and their respective source is available upon request.



(a) Employment (in full-time equivalents)

(b) Energy expenses (in thousands 2015 Euros)

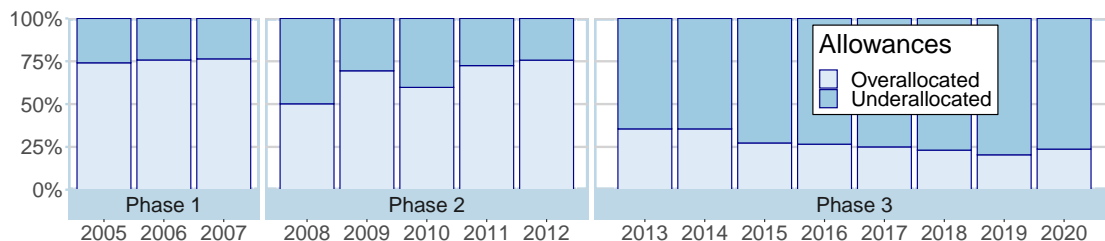
**Figure 3:** Averages of selected variables over time.

Note: Cohorts 1-3 consist of firms first regulated in Phases 1-3, respectively. The Matched control group consists of unregulated firms that are matched to regulated firms. The vertical axes are on a log scale.



One can see that firms regulated in 2005 are by far the largest energy consumers. This makes sense, since large emitters were regulated first. Note that these are energy expenses and that energy prices are responsible for some of the time variation. In terms of employment, treated firms seem to be more similar across phases, with now the second cohort being the largest. One can also see that regulated firms are far more energy-intense than unregulated firms, but are more comparable in terms of employment, even though there still remains a gap. It is difficult to hypothesize on the estimated treatment effect, as panel (a) does not show clear kinks at the treatment dates.

Another form of heterogeneity lies in the treatment stringency that a firm experiences from the regulation at different moments in time. As almost all allowances were handed out for free in the first two phases, one could argue that regulation was not stringent in these phases. It was also not uncommon that firms were over-allocated with free allowances, causing these firms to be effectively net subsidized by the ETS. In theory the allocation of allowances should not influence the decision of the firm, as the allocation does not influence the firm's opportunity costs. However, these allocations are likely to have mattered in practice, as firms might have not regarded the regulation as binding and might subsequently not have adjusted their behavior. Profits and financial constraints are mostly unaffected when all emissions are covered by freely allocated allowances.



**Figure 4:** *EU ETS stringency.*

*Note: The share of Dutch regulated firms that receive more (or less) allowances for free than their verified emissions. Source: authors calculations based on EUTL data accessed through [EUETS.INFO](http://EUETS.INFO).*

In phase 3, however, allocation mechanisms changed to auctioning as the default option. As many firms were exempted from the switch to auctioning, treatment stringency became more heterogeneous between firms, with some still receiving more allowances than needed, but most now receiving fewer than needed. [Figure 4](#) presents an overview of ETS stringency over time. Together with the changes in prices, shown in [Figure 2](#), this creates significant heterogeneity in the policy stringency both over time and across firms.

## 4 Methodology

### 4.1 Measures of competitiveness and investment intensity

We are interested in the ETS's effects on (1) competitiveness and (2) investments. We measure these concepts with three dependent variables, namely, (1a) the firm's employment and (1b) its profit margins, to inform us on competitiveness, and (2) its investment intensity, to inform us on investment outcomes. Tracking employment outcomes also allows us to evaluate whether domestic environmental regulation indeed led to job losses at home, an often heard counterargument to unilateral environmental policy. Profit margins directly evaluate the profitability of the regulated firms. They also show to what extent regulated firms were able to charge a price that was above their marginal costs, thus they also show if regulated firms were able to pass on additional costs of the regulation to their consumers. This ability likely decreases with the level of competition from abroad. Investment intensity estimates in how far firms are incentivised to invest into new technologies as a response to the regulation.

We measure employment in full-time equivalents (FTE) and use two definitions for the profit margin, (1) the gross profit margin and (2) the EBITDA margin. Gross profits measure the difference between turnover and the costs of goods sold. EBITDA measures earnings before interest, taxes, depreciation and amortization. We scale both measures by turnover to transform them into a margin. Gross profits are generally larger than EBITDA. They are also not influenced by a firm's financial operations, and thereby for example exclude a firm's income from holding activities.

We use investments into all fixed assets and scale them by turnover in order to scale by the size of the firm's activities. As an alternative specification we also scale investments by employment.

In order to further improve comparability across industries, we normalize the profit margins and investment ratios. We do this according to the following normalization:

$$\bar{x}_{jit} = \frac{x_{jit} - \mu_{it}^x}{\sigma_{it}^x}, \quad (4.1)$$

where  $j$ ,  $i$ ,  $t$  refer to firm, industry and time, and  $\mu^x$  and  $\sigma^x$  are  $x$ 's industry-time mean and standard deviation. Note the different subscripts, indicating that the variable is normalized using the industry-specific mean and standard deviation. This way the resulting transformed variable is comparable across industries.

If variables are in monetary terms, they are deflated such that they can be compared over time. For this deflation we use Eurostat's industry producer price index for the Netherlands.

## 4.2 General identification strategy

To identify the effects of the ETS, we use the fact that not all manufacturing firms in the Netherlands are regulated under the ETS. Regulation is on the plant level and there are mainly two criteria for inclusion in the ETS, either (1) through exceeding a certain sector-specific threshold related to energy input or production capacity, or (2) through incorporating specific processes that imply automatic regulation.<sup>6</sup> This implies that one can attempt to find comparable control firms for each treated firm that are both active in comparable production processes and are comparable in terms of size, employment characteristics and energy input.

To then evaluate the causal effects of the EU ETS, we apply two empirical policy evaluation methods. The key in these methods is to use detailed microdata on observed firms to compare the outcomes of treated firms, i.e. firms receiving regulation, to the outcomes of *comparable* control firms, i.e. firms not regulated (yet). The first method is a matched two-way fixed effects regression and the second method is a less parametric DiD design suggested by Callaway and Sant’Anna (2021).

In general two main steps can be identified in this evaluation process, namely (1) matching or weighting, in which we score firms across treatment status based on their similarity, and (2) comparison, in which we either regress our outcome variable on treatment status or take differences in outcome variables across treatment status. The second step utilizes the weights established in the first step. In both methods, we estimate a treatment effect for each cohort-phase combination, thus controlling for the heterogeneity between treated cohorts and in different phases.

## 4.3 Sample selection

One estimation issue arises from firm exit and, to a lesser degree, firm entry, from and to the sample. As we are dealing with anonymized microdata it is not possible to determine if such an exit is due to closure of the firm, an acquisition by another firm or due to changes in the firm structure. To minimize the effect that sample composition could have on our results, we curtail our sample to firms that we observe continuously from two years before to three years after treatment start. Unregulated firms also face this requirement when considered for matching.

We also enforce a common support for all of our covariates (employment, energy costs, turnover and total wage bill) between treated and control groups in the baseline years. The baseline year is two years before treatment start, to allow for

---

<sup>6</sup>For a detailed overview see Annex I of European Parliament, Council of the European Union (2003).

one year of treatment anticipation.<sup>7</sup>

## 4.4 Matched TWFE method

We break the matching and regression up in the following two subsections. The first one explains the matching that provides the weights, and the second one presents the details of the regression.

### 4.4.1 Matching

The goal of matching is to select similar observations across treatment status from the data. In general a matching algorithm provides a similarity score between each pair of observations in the sample data. If provided with  $n$  observations, the matching outcome matrix  $M$  has dimensions  $n \times n$ . For our TWFE application the pair information is dropped and only those observations with a high enough similarity score to any other observation across treatment status are kept, collapsing the matching information from  $M$  to a binary  $n \times 1$  vector, indicating for each firm if it will be kept in the estimation or not. Observations in the non-treated group that do not have a high enough similarity score with a treated observation are thus dropped from our sample. This way matching boils down to sample selection.

The matching outcomes are used to select the sample for our TWFE regression. All observations are kept of firms that are matched, either in the treatment group or the control group (i.e. have a value of 1 in the  $n \times 1$  vector). This effectively is a special form of weighting, as the weights are either 1 (for the matched) or 0 (for the non-matched).

We base our matching on all variables that affect the probability of treatment and try to align it with other studies. We only match within the two-digit industry code, and base the similarity on a firm's employment, energy costs, turnover and total wage bill as well as the squared values of these variables. Matching happens two years before treatment start, to account for anticipation. Our matching algorithm is further elaborated in [Algorithm 1](#) in [Appendix C](#).

---

<sup>7</sup>Before phase 1 the important directive for the establishment of an ETS was passed in 2003, before phase 2 the national allocation plans had to be published in 2006, and before phase 3 the commission passed directive 2011/540/EU in 2011, extending the scope of regulated greenhouse gases and industries.

#### 4.4.2 TWFE regression

Using the resulting matched sample, we can estimate the impact of the EU ETS's phases on each cohort's outcomes. Our two-way fixed effects regression looks as

$$y_{jt} = \sum_{c \in C} \sum_{p \in P} ETS_j^c \times P_t^p \times \mathbb{1}_{\{p \geq c\}} \alpha^{cp} + \gamma_j + \gamma_t + \varepsilon_{jt} \quad (4.2)$$

where  $y$  is the outcome variable of interest and subscripts  $j, t$  refer to the firm and year.  $ETS^c$  is a dummy variable that is equal to one if firm  $j$  is in cohort  $c$ .  $P$  is a dummy that is equal to one if year  $t$  is in ETS Phase  $p$ . As there are three phases in our data range, we have  $C, P \in \{1, 2, 3\}$ . The interactions of the two variables thus present the treatment indicators of our DiD regression. The coefficients of interest are the corresponding  $\alpha$ s, with one coefficient for each of the six post-treatment cohort-phase combinations (i.e. cohort 1-phase 1 through cohort 3-phase 3).

We include firm and optionally year fixed effects, but abstain from including time varying controls, as these are likely to be affected by the treatment itself.  $\varepsilon$  is the error term, which we allow to be heteroskedastic and serially correlated. We estimate the model using ordinary least squares.

The dependent variable  $y$  is either the log-transformed number of employees or a normalized EBITDA margin when interested in the competitiveness effects from the EU ETS, and investments in fixed assets, scaled by turnover and normalized, when interested in the investments response.

### 4.5 CS2021's DiD estimator

Our second approach also relies on matching and a DiD design, but it does so in a less parametric fashion. It follows the approach by Callaway and Sant'Anna (2021). We briefly motivate the use of another estimator by discussing the drawbacks of TWFE, before discussing the estimator itself.

#### 4.5.1 Potential problems with TWFE in staggered DiD

Recent econometric literature has pointed out several problems with TWFE estimations in DiD settings like ours. This literature focuses on the potential biases in TWFE estimators applied to settings with staggered treatment adaption and potentially heterogeneous treatment effects (see e.g. Daw & Hatfield, 2018; de Chaisemartin & D'Haultfoeuille, 2022; Goodman-Bacon, 2021). This is exactly the case in our setting, in which firms get treated in different phases and in which we both expect these different cohorts to react differently and in which we assume the effect to be time (or phase) dependent.

The key problem of TWFE in such cases is that the derived estimator for the Average Treatment Effect on the Treated is a weighted average over the ATTs of the different treatment groups at different times, without explicitly appreciating this and without being able to control the weights of the individual ATTs. The estimator thus only gives a clearly interpretable ATT if treatment effects are constant both over time and between treatment groups.

As described in detail by Goodman-Bacon (2021), these weights can in some cases be negative, resulting in a distorted estimate. Even if the weights are non-negative, they are unobserved and mostly determined by the group size underlying the estimate and the distance to the start of the treatment.

Another problem of the matched-TWFE estimator is that most of the matching information is lost in the regression step. Matching is purely used for sample selection, while the link between matched treated and non-treated units is not taken into account in the estimation. This means that a control firm that is matched to a treated firm in cohort 1 will serve as a control also for treated firms in cohorts 2 and 3, and so on.

#### 4.5.2 The estimator

To address the above mentioned issues, we make use of the estimator developed in Callaway and Sant’Anna (2021). Its main advantage lies in the fact that it estimates ATTs for each treatment cohort – the group of firms starting treatment in the same phase – and at each year into the treatment. It also allows for different aggregations of those estimates, enabling us to restrict the type of heterogeneity.

The estimator is in essence an application of the doubly-robust DiD estimator of Sant’Anna and Zhao (2020) to staggered settings. It pays close attention to the conditioning on covariates, combining both inverse probability weighting (see Abadie, 2005) as well as outcome regression adjustment (see Heckman et al., 1997). The latter is also frequently used in adjusted versions in comparable ETS papers like Martin et al. (2014) or Löschel et al. (2019).

While the inverse probability weighting tries to re-balance the control group based on their probability of being treated, thus in fact on their similarity to the treatment group, the outcome regression adjustment tries to take out trends in the outcome variable that are covariate dependent. The CS estimator is therefore consistent as long as the covariate conditioning is correctly specified by either one (or both) of the two covariate conditioning strategies (therefore referred to as being “doubly-robust”).

An additional advantage of this weighting is that each treated firm is linked to a specific set of control firms and is only compared to these control firms. This is in contrast with the matched TWFE estimator.

The estimator for each cohort,  $c$ , and year,  $t$ , is then a common average treat-

ment effect DiD estimator. It compares the outcome of each firm in year  $t$  to the firm’s own outcome in the base year,  $b$ , and to that of the weighted average difference in outcomes between  $t$  and  $b$  of the respective control group for this firm. For this weighting, both inverse probability weighting and the outcome regression adjustment are used.

The following equation specifies the estimated ATT for cohort  $c$  and year  $t$

$$\hat{\alpha}_{ct} = \frac{1}{N} \sum_{j \in \mathcal{J}} \left[ \underbrace{(\hat{w}_{jc}^{treated} - \hat{w}_{jc}^{control})}_{\text{Inv. prob. weight.}} (y_{jt} - y_{jb} - \underbrace{\hat{m}_{jct}(X_j, \hat{\lambda}_{ct})}_{\text{Outcome reg.}}) \right], \quad (4.3)$$

with  $N$  the number of firms and  $\mathcal{J}$  the set of all firms,  $y_{jt}$  the dependent variable,  $X$  as pre-treatment controls, and  $j, c, t$  referring to firm, cohort and year. *treated* and *control* refer to the treatment status, i.e. regulated or control firms.

$\hat{w}_{jg}^{treated}$  and  $\hat{w}_{jg}^{control}$  are the weights that adjust for the probability of being treated. They are 0 if a firm is not in the respective group and give higher weights to control firms that are more similar to the treated firm, given a set of covariates.  $\hat{m}_{jct}(X, \hat{\lambda}_{ct})$  represents the bias adjustment from an outcome regression, thus deducting the predicted development of  $y$  based on  $X$ , under the assumption that the firm had not been treated. More information on both adjustments and their exact definition can be found in [Appendix B](#).

In this setting, we cannot enforce matching within an industry and thus include industry dummies for the three sectors containing at least one firm regulated in all phases in  $X$ .<sup>8</sup> Employment, turnover, wages and energy expenses as well as their squared values enter as predictors for inverse probability weights and the outcome regression as well. As in the TWFE setting, we assume one year of anticipation, pinning down the base year at two years before the treatment starts. The corresponding standard errors are bootstrapped and clustered at the firm level.

Two sets of candidate control firms can be considered, namely (1) the entire population of firms that has not been treated up to  $t$ , or (2) only the set of firms that will never be treated. For our main specification, we choose to use all not-yet-treated firms as controls, since these will likely be more similar to earlier treated firms. Results for the never-treated control group are presented in the discussion section.

There is no guarantee that this set of control firms exhibits parallel trends in absence of the ETS. To test for parallel pre-treatment trends, we employ a placebo test. By testing whether pre-treatment ATTs (always assuming that the base year is one year before  $t$ ) are different from zero, the test indicates whether

---

<sup>8</sup>These are manufacture of food products and beverages, manufacture of chemicals and chemical products, and manufacture of other non-metallic mineral products. Adjusting the choice of these sectors barely affects our results.

any disparities between the (weighted) treated and control units occurs during the pre-treatment years. We once use a Wald test to for joint significance of these placebo tests and once apply the aggregation outlined in the next paragraph and compute a respective confidence interval to check if it contains zero or not.

Equation 4.3 presents ATT estimations for each cohort-year pair. This allows for more heterogeneity than we assume to be present, as we are interested in cohort-phase effects. Following Callaway and Sant’Anna (2021), we aggregate the separate cohort-year ATTs to a respective cohort-phase,  $\tilde{c}$  and  $\tilde{p}$ , aggregate using cohort-year weights,  $v^{\tilde{c},\tilde{p}}(c,t)$ , defined as:

$$v^{\tilde{c},\tilde{p}}(c,t) = P[t|c = \tilde{c} \text{ and } t \in \tilde{p}] \mathbb{1}_{\{c=\tilde{c}\}} \mathbb{1}_{\{t \in \tilde{p}\}}. \quad (4.4)$$

All weights are non-negative and add up to one within each cohort-phase. These weights are then used in the aggregation

$$\hat{\theta}^{\tilde{c},\tilde{p}} = \sum_{c \in \{1,2,3\}} \sum_{t=2005}^{2020} \hat{v}^{\tilde{c},\tilde{p}}(c,t) \hat{\alpha}_{ct}, \quad (4.5)$$

in which  $\theta$  is the cohort-phase aggregated ATT.

For inference, a bootstrap algorithm calculates a 95% confidence interval around each estimator  $\hat{\theta}^{\tilde{c},\tilde{p}}$ . The algorithm repeatedly draws a subsample from the original sample, keeping the sizes of the cohorts proportional, and estimates the  $\hat{\alpha}_{c,t}$ s and the  $\hat{\theta}^{\tilde{c},\tilde{p}}$ s. From this distribution of estimators the 2.5th and 97.5th percentile determine the 95% confidence interval.

## 5 Findings

This section presents the findings of the proposed methodologies from the previous section. First we discuss the TWFE results together with the associated matching outcomes. We use this opportunity to also discuss the differences between treated and control firms, as the control group is clearer defined in the TWFE than in the CS setting. We then present the results from the CS estimation and subsequently discuss some differences in the results.

### 5.1 TWFE

#### 5.1.1 Matching outcomes

As described in the matching algorithm, we require firms to be observed around the treatment start. We only keep firms that are observed all years from 2 years



before to 3 years after treatment start. This leaves 115 ETS firms in our sample. Of these, 105 are matched to control firms.<sup>9</sup>

Table 1 presents a balancing table, showing the means of the matched firms for the three cohorts. The balancing table shows that regulated firms are larger and invest (relatively) more than unregulated firms. Profit margins are rather comparable between treated and control firms. Even though the matching selects similar firms, some of the size difference remains. This is inherent as the ETS by design regulates larger and more energy-intensive firms. Remaining level differences therefore need to be accounted for within the estimation. The control group of firms in the CS setting will be different and varies each year, as the weights are also based on the outcome regression adjustment. They will, however, be tailored even more to the individual firms thus reducing the size differences.

Further matching outcomes are presented in Appendix C. There Figure C.1 through C.3 provide distribution plots for the matching variables both before and after matching. This is mostly because firms that are very dissimilar to regulated firms are excluded. These figures show that matching improves the comparability of the distributions across treatment status.

### 5.1.2 Regression results

Table 2 presents the results of our TWFE estimation from Equation 4.2 for employment, the gross profit margin and the investment ratio as dependent variables. The regressions include firm fixed effects and either phase or year fixed effects. Cohort and industry instead of firm fixed effects do not lead to substantially different results.

Most of the estimated coefficients are statistically insignificant, indicating that we do not establish many side-effects from ETS regulation in this estimation. The significant estimates, however, have the predicted sign and indicate a reduction in employment for cohort 1 firms. The effect seems to be persistent but is most clearly estimated right at the beginning of the treatment start. On the one hand these firms are the most energy-intensive and are thus most likely to be affected by the regulation, making it reasonable that these firms had to reduce their employment the most as a response to the regulation. On the other hand, the early EU ETS phases are widely deemed to lack stringency, making the result surprising. Cohort 2 shows a similar effect in phase 2, but the estimate is not statistically significant.

The estimated employment effects are also quite sizable, as cohort 1 experiences a nearly 8% decrease in phase 1 compared to pre-treatment years. This effect is mostly maintained in phase 2, as phase 2 employment is 9% lower compared

---

<sup>9</sup>Doing the estimation on a fully balanced panel greatly reduces the power of the estimation, but the qualitative results remain similar.

**Table 1:** *Balancing table for the variables used in the analysis. The cohorts refer to the EU ETS firms' first regulated phase. Treated and control refer to matched EU ETS and non-ETS firms. The groups presented here are the ones in the TWFE estimation. The treated group is almost identical in the CS estimation. Values correspond to the group means two years before treatment start. Monetary variables are in millions of 2015 Euros (Investment/Employment 2015 TEuros). Standard deviations are in brackets.*

	Cohort 1			Cohort 2			Cohort 3		
	Control	Treated	Difference	Control	Treated	Difference	Control	Treated	Difference
Employment	253,78 (242,27)	469,61 (372,46)	215,83 (52,87)	246,86 (226,40)	928,62 (1311,72)	681,76 (225,45)	244,43 (211,32)	276,20 (269,27)	31,76 (68,79)
Gross Profit Margin	0,47 (0,16)	0,51 (0,19)	0,04 (0,03)	0,43 (0,18)	0,57 (0,16)	0,13 (0,03)	0,44 (0,19)	0,45 (0,26)	0,02 (0,07)
Investment/Turnover	0,04 (0,08)	0,06 (0,10)	0,02 (0,01)	0,05 (0,10)	0,05 (0,05)	0,00 (0,01)	0,03 (0,09)	0,06 (0,08)	0,03 (0,02)
Energy costs	2,01 (5,14)	11,98 (14,19)	9,97 (1,95)	2,87 (6,36)	8,24 (11,13)	5,37 (1,95)	2,25 (3,50)	7,09 (9,15)	4,84 (2,30)
Wages	9,53 (9,51)	21,09 (19,71)	11,56 (2,74)	9,49 (8,96)	38,32 (52,65)	28,83 (9,05)	10,80 (10,62)	12,39 (12,16)	1,59 (3,12)
Turnover	84,85 (94,17)	236,19 (329,58)	151,34 (44,90)	97,02 (97,01)	346,64 (566,72)	249,63 (97,40)	117,37 (159,51)	183,15 (226,12)	65,78 (57,53)
EBITDA Margin	0,10 (0,09)	0,12 (0,10)	0,02 (0,01)	0,09 (0,11)	0,12 (0,12)	0,04 (0,02)	0,08 (0,10)	0,12 (0,08)	0,04 (0,02)
Investment/Employment	10,33 (13,96)	26,38 (57,03)	16,06 (7,82)	14,81 (25,78)	15,06 (13,58)	0,25 (2,88)	13,04 (29,34)	39,95 (59,33)	26,91 (14,97)
Observations	215	55		231	34		223	16	

to pre-treatment years. Cohort 1’s phase 3 estimate is similar but statistically insignificant.

The estimates for the other outcome variables show little effect of the regulation. We can neither establish a statistically significant effect on profitability or investments in this estimation.

**Table 2:** *Results from the matched-TWFE regression.*

	Employment		Profit Margin		Investment Ratio	
Cohort 1 × Phase 1	-0.078** (0.033)	-0.077** (0.033)	0.002 (0.046)	0.001 (0.046)	-0.002 (0.052)	-0.002 (0.052)
Cohort 1 × Phase 2	-0.088* (0.050)	-0.088* (0.050)	0.026 (0.098)	0.025 (0.098)	0.009 (0.047)	0.010 (0.047)
Cohort 1 × Phase 3	-0.092 (0.074)	-0.094 (0.073)	-0.046 (0.170)	-0.047 (0.170)	0.003 (0.050)	0.004 (0.050)
Cohort 2 × Phase 2	-0.048 (0.058)	-0.047 (0.058)	-0.095 (0.104)	-0.097 (0.104)	0.067 (0.063)	0.068 (0.063)
Cohort 2 × Phase 3	0.008 (0.070)	0.005 (0.070)	-0.133 (0.152)	-0.136 (0.152)	-0.014 (0.069)	-0.013 (0.068)
Cohort 3 × Phase 3	0.101 (0.094)	0.096 (0.094)	0.232 (0.175)	0.233 (0.174)	-0.114 (0.075)	-0.113 (0.075)
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Phase FEs	Yes	No	Yes	No	Yes	No
Year FEs	No	Yes	No	Yes	No	Yes
Observations	6,273	6,273	6,249	6,249	6,185	6,185
Adjusted R2	0.902	0.903	0.655	0.655	0.114	0.114

The dependent variables are the log of the number of employees, the gross profit margin, and the investment to turnover ratio. The margin and the ratio are normalized on the industry level as in (4.1). Standard errors are clustered at the firm level and reported in brackets. Stars refer to \*:  $p < 0.10$ , \*\*:  $p < 0.05$ , \*\*\*:  $p < 0.01$ .

## 5.2 CS estimation

The cohort-phase estimates for the less parametric estimation from Equation 4.5 are presented in Table 3. For employment, the conclusions differ somewhat. Cohort 1 firms still reduce employment in phase 1, even though the effect is now only significant at the 10% level, but for phase 2 and 3 there are no statistically significant effects from treatment anymore. Cohort 2 phase 2’s estimate is now statistically significant, indicating a reduction in employment of about 7%. For the other cohort-phase combinations the estimates are statistically insignificant, as before.

**Table 3:** *Results from the less parametric DiD method.*

	Employment	GP margin	Investment ratio
Cohort 1 $\times$ Phase 1	-0.058* (-0.122,0.003)	0.119 (-0.262,0.453)	-0.318** (-0.508,-0.090)
Cohort 1 $\times$ Phase 2	0.034 (-0.075,0.111)	-0.187 (-0.528,0.533)	-0.466** (-0.733,-0.142)
Cohort 1 $\times$ Phase 3	0.001* (-0.106,0.074)	0.090 (-0.351,0.845)	-0.325** (-0.517,0.006)
Cohort 2 $\times$ Phase 2	-0.070 (-0.123,-0.021)	-0.005 (-0.192,0.260)	0.036 (-0.108,0.165)
Cohort 2 $\times$ Phase 3	0.002 (-0.209,0.076)	0.057 (-0.217,0.266)	-0.012 (-0.118,0.098)
Cohort 3 $\times$ Phase 3	0.069* (-0.158,0.342)	0.449*** (0.082,1.430)	-0.198* (-0.454,0.005)
Number of firms	1012	1012	1012
Regulated firms.	113	113	113
Pre-treatment	0.001 (-0.010,0.011)	-0.004 (-0.047,0.047)	0.013* (-0.003,0.049)

The dependent variables are the log of the number of employees, the gross profit margin, and the investment to turnover ratio. The margin and the ratio are normalized on the industry level as in (4.1). Bootstrapped 95% confidence intervals are in brackets. The pre-treatment estimate pools all placebo pre-treatment estimates. Stars indicate if 0 contained in confidence interval: \*: 90%, \*\*: 95%, \*\*\*: 99%

Further, cohort 3 firms increase profit margins in phase 3 with 0.45 industry-level standard deviations. The other cohorts experience no changes in their profit margins due to ETS regulation.

For investments, the effects are predominantly negative. Cohort 1 (cohort 3) firms reduce their investments-to-turnover ratio by around 0.3 (0.2) industry-level standard deviations, while cohort 2 firms do not statistically significantly respond to ETS regulation with their investment ratio. This is the sharpest deviation between the results of the TWFE and the CS estimates. Reductions in investments are inconsistent with the Porter hypothesis, but together with the negative effects on employment, could indicate a downsizing of EU activities in these firms.

### 5.3 Differences between TWFE and CS

The main differences between the results of the two estimators lie in the estimates of the effect on investments for cohort 1 and in the employment effect for cohort 1 in phases 2 and 3. Besides this, most coefficients align and mostly vary in terms of uncertainty.

The most likely reason for this deviation lies in the difference of the control groups that underlie both estimations. As outlined above, one problem with the TWFE estimator is that a treated cohort's outcomes are compared to every control firm's outcome. The matching links are lost in the regression step and thus the comparison is not between the most similar firms. The CS method compares cohorts with their own specific control firms. In the TWFE setting cohort 1-phase 1 observations could be compared to control firms that were matched to cohort 3. As cohort 1 and 3 differ, their matches likely also differ, undermining the quality of the comparison.

To make sure that the CS results are indeed based on a more specific and more comparable control group, the following section we will estimate different specifications in which we vary the underlying observations of the CS estimator.

## 6 Discussion

This section presents a discussion on the underlying assumptions and tests the robustness of our results to violations of these. By doing so, we also test how robust our results are to changes in the underlying control group. We also present results for slightly different dependent variables, and discuss our relation to the results in the literature. The presented results will in most cases be purely based on the CS estimation, as this is our preferred specification, but the TWFE results do not give additional insights above the presented and discussed ones.

## 6.1 Parallel trends

All DiD estimates rely on the parallel trends assumption to hold. Even though there is no formal test for parallel trends, one can perform some checks. We run placebo tests to see if pre-treatment periods experience differences across treatment and control groups. If so, there would be some evidence that the trends of the treatment and (matched) control would have not run in parallel in absence of the treatment.

For the TWFE method, we bring the treatment date 1, 2 and 3 years forward, drop all estimates after the actual treatment date, and see if the treatment coefficients result are statistically significant. [Figure 5](#) presents the results of these tests. For none of the date-shifts the placebo test results in statistically significant findings, implying that we cannot reject the null hypothesis of parallel trends.

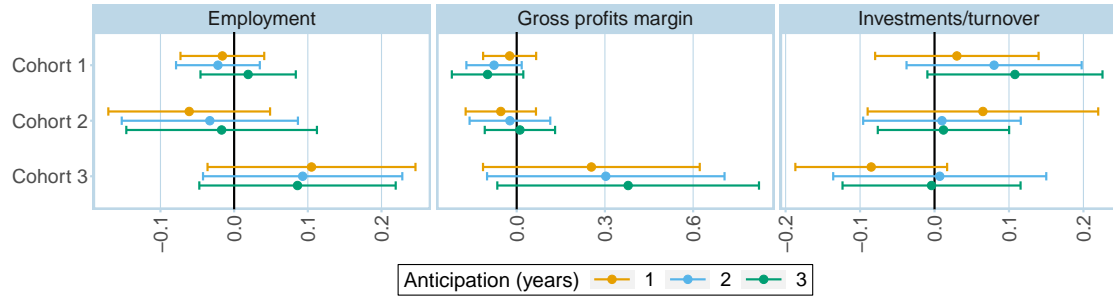
For the less parametric DiD method, we perform two similar checks. The ATTs can be estimated individually for each of the pre-treatment years, by always choosing the baseline to be the year before. We then use a Wald test to test the joint significance of these estimates; we do not reject the null for any of the three variables. We also precede in our cohort-phase setup, by aggregating all pre-treatment estimates and constructing a confidence interval in the same way as for our aggregated estimates. These results can be found in the first row of [Table 3](#). They always contain the zero in the 95% confidence interval as well.

The parallel trends tests were rejected for the analysis with the EBITDA margin. Therefore the main analysis focuses on the related gross profit margin, for which the parallel trends assumption is not rejected.

To test the robustness of the results, we adapt the matching to make the assumption more likely to hold. We have done this by using the trends, along-side the regular values, of the matching variables in the matching steps, both for the TWFE method as well as for the CS method. The resulting estimates can be found in [Figure 6](#). The results for investment completely align; for employment the coefficients are now more negative, especially for cohort 3, and similarly gross profits turn negative for cohort 3 and cohort 2-phase 2. The overall picture, however, remains the same, indicating a reduction in investments for cohort 1 and a reduction in employment for cohorts 1 and 2. The results for cohort 3 (the smallest cohort) seem to be rather volatile.

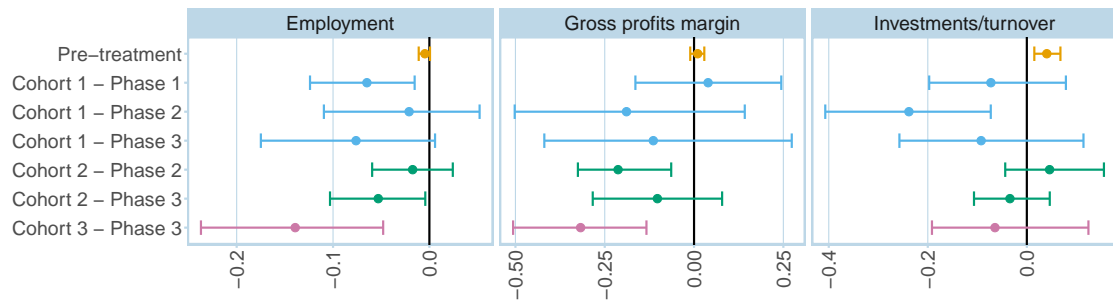
## 6.2 SUTVA

The Stable Unit Treatment Value Assumption (SUTVA) is the second necessary assumption to identification in a DiD setting. It in essence implies no spillovers between firms across treatment status. As our analysis is on the firm level, a large source of spillovers, namely that between plants in the same firm, is accounted for



**Figure 5:** *Placebo tests for parallel trends in the TWFE model.*

*Note: Placebo tests for testing the pre-treatment trend for each cohort. The colors refer to the number of years that the treatment is brought forward. E.g. 1 year of anticipation for cohort 1 tests whether there is a treatment effect in 2004. Whiskers indicate 95% confidence intervals.*



**Figure 6:** *EU ETS effects when matching on pre-trend.*

*Note: Cohort-phase coefficient estimates when the trend of the dependent variable is included in covariates to strengthen common trends. Whiskers indicate bootstrapped 95% confidence intervals. Colors refer to the different cohorts.*

already.

As regulation is on the plant level and our outcome variables are on the firm level, not all activities of the firm are regulated. This likely biases the estimates towards zero. Note that this is not a drawback per se, as this is simply how the ETS functions. The estimates are accurately representing the effect of ETS regulation on firm performance. If interested in the question what happens if all emissions were regulated, the estimates are biased towards zero.

On the other hand, there might be positive spillovers between firms through the markets that they operate in. As competition is relative, one firm's hardship can be another firm's opportunity. This can both be on the output market, as unregulated firms obtain a relative cost advantage from not being regulated, as well as on the input markets, as potential downsizing of regulated firms allows unregulated firms to snatch up employees or input supplies. This would inflate the estimates.

As both these biases to some extent relate to treatment stringency, one could compare the estimates for the different phases in the data. From [Figure 2](#) and [Figure 4](#) it becomes clear that later phases are more costly to regulated firms. [Figure 1](#) also shows that more installations of the same owner are regulated in phase 3. Estimates for the later phases should therefore suffer less from the bias towards zero, as more emissions of the firm are regulated, and more from the bias away from zero, as the relative disadvantage from regulation is exacerbated. If these biases exist, in both cases they should result in larger estimates, in absolute terms, for the later phases. [Table 3](#) does not provide evidence for either bias, as estimates for later phases within the same cohort are not further away from zero.

### 6.2.1 Treatment anticipation

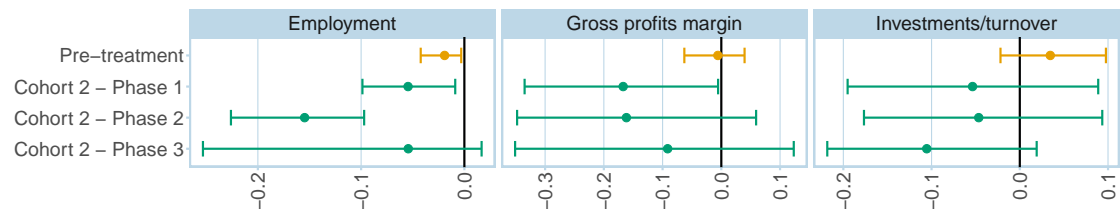
As explained in [section 3](#), the main reason for why there are so many firms that are only regulated in the second phase is that the Dutch government excluded many firms from regulation in the first phase. As these exemptions are public information, it seems likely that these firms expected to be regulated in phase 2. If so, the firms in cohort 2 would have already anticipated treatment in 2003, which would violate our assumptions on the anticipation, and would also make these firms an improper control for cohort 1 in phase 1.

On the other hand, this might enable us to roughly disentangle an anticipation from an actual treatment effect, by treating cohort 2 as already being regulated in phase 1. The estimate from such an experiment also provides cohort 2's anticipation effect of being regulated in phase 1, and provides an adjusted estimate of the effect in phases 2 and 3, with an adjusted control group and base year (2003).

The results for this estimation can be found in [Figure 7](#). The results somewhat align with the main results for phases 2 and 3, although the statistical significance



differs. Interestingly the negative employment effect in phase 2 was preceded by a small negative effect in phase 1 already, which could imply that firms did anticipate the regulation and started reducing their employment in anticipation. The same can be concluded for the gross profit margin.



**Figure 7:** *EU ETS effects when treating cohort 2 as of Phase 1.*

*Note: Cohort-phase coefficient estimates for cohort 1 when allowing it to show a regulation response as of Phase 1. Whiskers indicate bootstrapped 95% confidence intervals. Colors refer to the different cohorts.*

## 6.2.2 Compare to never-treated firms

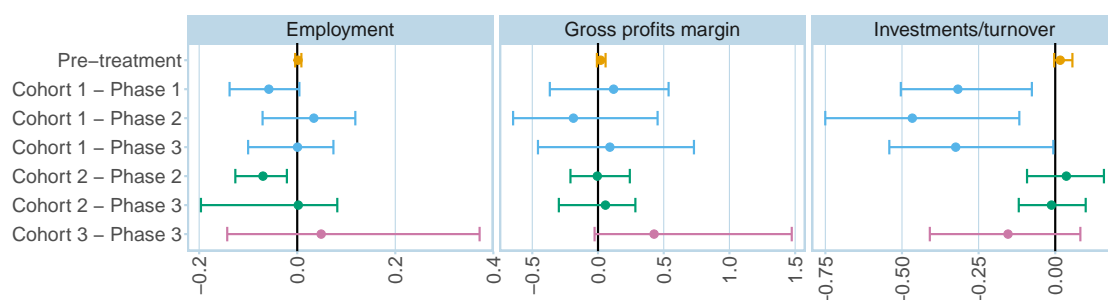
To further alleviate concerns about potential anticipation effects in the control group, we can rid the control group of firms that in later phases become regulated. The control group then exclusively consists of firms that are never regulated under the EU ETS. The advantage is that the control group is then less likely to contain anticipation effects to upcoming regulation, as was indicated in the previous subsection. The disadvantage is that restricting the set of available controls makes the matches poorer.

Figure 8 shows the results of this exercise. Most notable is the close similarity of the results to our main analysis, making us further confident in our estimates.

## 6.3 Alternative measures as dependent variables

To test if our results rely on our choice of the dependent variable, we test if (1) the scaling of investment and (2) the choice of the profit margin is important for our results. We thus first scale investments (still into all fixed assets) by the number of employees in full-time equivalents. The results are presented in Figure 9. The results are in line with the main specification, in which we scaled investments by turnover.

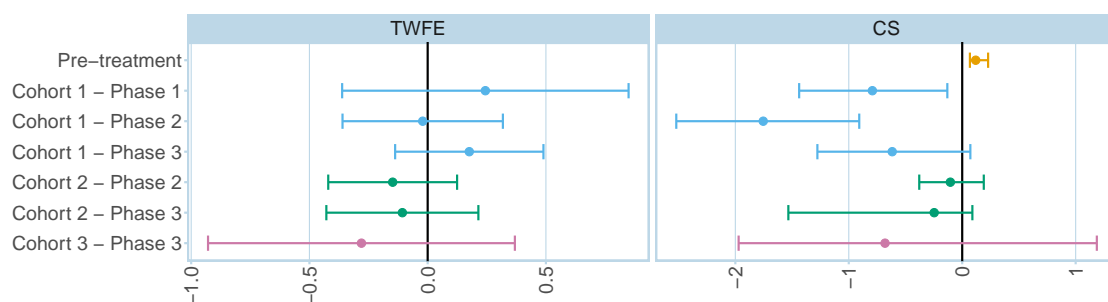
Second, we look at the EBITDA margin instead of the gross profit margin. EBITDA is often used as a more harmonized measure of earnings, as it takes the gross profits before the financial result. As mentioned earlier the EBITDA margin does not pass the placebo tests and hence we have rely on the gross profit margin



**Figure 8:** *EU ETS effects when strictly comparing to never treated control firms.*

*Note: Cohort-phase coefficient estimates for cohorts 1-3 when strictly allowing for comparisons with never-treated control firms. Whiskers indicate bootstrapped 95% confidence intervals. Colors refer to the different cohorts.*

instead. For completeness the results for the EBITDA margin can be found in [Appendix D, Figure D.4](#), and have similar signs as that of the gross profit margin, with, however, smaller confidence intervals.



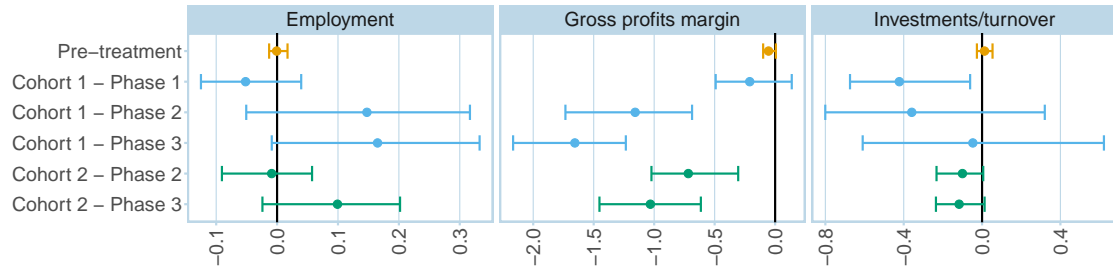
**Figure 9:** *EU ETS effects on investments over employment.*

*Note: Cohort-phase coefficient estimates for cohorts 1-3 for investments scaled by the number of employees (FTE). Both the TWFE and CS results are provided. Whiskers indicate 95% confidence intervals, which for CS are bootstrapped. Colors refer to the different cohorts.*

## 6.4 Fully balanced panel

To see in how far our results are sensitive to firms exiting our sample at a later stage, we redo the estimation on a fully balanced panel. This greatly reduces the number of observations and firms, which also prevents us from reporting cohort 3 findings due to privacy requirements from CBS.

The results are presented in Figure 10 and show that while the results for investments are still in line with the main results, those for the profit margin are quite different. The results indicate that on this smaller panel, the regulated firms experience large losses in profitability. As this is based on few observations, we assume that these results are based on few outliers, but we take this as further evidence that the profitability results are rather volatile and inconclusive. For employment our results confirm the early reduction in employment in phase 1, but also show that this effect was likely temporary.



**Figure 10:** *EU ETS effects when enforcing a balanced panel.*

*Note: Cohort-phase coefficient estimates for cohorts 1-3 when enforcing a balanced panel, with thus only 135 underlying firms (52 treated). Each firm in the sample is observed for all the years in the sample. Whiskers indicate bootstrapped 95% confidence intervals. Colors refer to the different cohorts.*

## 6.5 Comparison to literature

Most articles that have studied the ETS's effect have found little side effects from the regulation. In a sense our study's overall findings are not too different from the literature's findings. We do not find conclusive indications of a reduction in profit margins and the findings of negative employment effects seem temporary. However, we do find that cohort 1 reduces investments in all phases and employment in phase 1. We have shown that cohort 1 mostly consists of the most energy-intensive firms, which is likely linked to this result. Other studies do not reach these conclusions.

The reason for the negative effects could be based in the high export-orientation of the Dutch economy. The exposure to trade might have incentivized regulated firms to downscale their operations within the Netherlands as a response to the regulation, even before the regulation became more stringent. Such anticipation of future strictness of the regulation would also be in line with the anticipation effect that we find for cohort 2 firms. By disentangling the effects for the very energy-intensive firms in cohort 1 from the later, less energy-intensive cohorts, we show that heterogeneity between the cohorts plays an important role. Not all firms will

respond the same to the EU ETS. The fact that the negative employment effect is not maintained in the last phases, however, suggests that some effects might have been temporary.

Like Löschel et al. (2019) we also find statistically significant effects rather in the less-parametric DiD setting than in the TWFE regression, even though our results differ. This highlights the importance of choosing the right estimation methods when studying the ETS'effects.

## 7 Conclusion

This paper studies the effects of the EU ETS on the competitiveness and the investment behavior of regulated manufacturing firms in the Netherlands. Motivated by large differences in energy intensity between firms and by an increase in regulatory stringency over time, we pay special attention to treatment effect heterogeneity. We employ two difference-in-differences designs to estimate the average treatment effect on the treated. As a benchmark, we employ a classical matched two-way fixed effects regression. To better allow for heterogeneity, we also employ a recent, more flexible DiD method introduced by Callaway and Sant'Anna (2021).

We estimate the effects of EU ETS regulation on employment, the gross profit margin and investments over turnover. To capture the heterogeneities we estimate the effect of phases 1-3 in the EU ETS for each cohort, whereby a firm belongs to one of the three cohorts if it was first regulated in that respective phase. We make use of public data from the European Union Transaction Log and restricted-access microdata from Statistics Netherlands.

Our matched TWFE approach results in some evidence of employment loss in cohort 1 firms. For phase 1 and 2 these firms lose around 8% of employment compared to the non-regulated baseline. For profits and investments the TWFE method does not yield statistically significant findings.

In the preferred and more flexible DiD design, we find that cohort 1 firms reduce employment by 6%, but this finding is only significant at the 10% level and does not persist in later phases. Cohort 2 firms see a 7% decrease in employment in phase 2. For the other cohort-phase combinations the results are not statistically significant. The results with the profit margin as dependent variable are inconclusive. Cohort 1 firms reduce investments across all phases. The other cohorts do not seem to respond to regulation with changes in investments.

Taken together, our findings point to some employment loss in the Dutch manufacturing industry due to the EU ETS regulation. This is most notable in the first regulation phases for cohort 1 and 2. In phase 3 employment is restored to baseline levels for both cohorts. The temporary employment loss is between 7-9% in our main TWFE and flexible DiD estimations. A priori we expected such an

effect to be higher in later phases, when treatment stringency increased, but the result might indicate that firms reacted early, which is also consistent with some evidence for anticipation effects that we find by studying later treated firms.

We further find a negative effect on investments for cohort 1 firms, especially in phase 1 and 2, but only when applying the more flexible DiD method. These firms lose between 0.32 and 0.47 industry-level standard deviations in their investments to turnover ratio. We can conclude that a Porter Hypothesis style boost to investments in response to environmental regulation seems unlikely. For profitability the results are mixed and often statistically insignificant.

We thoroughly test the two main underlying assumptions of the DiD methodology, namely that of common trends and SUTVA. Common trends are tested for with pre-treatment placebo tests. We have only interpreted the findings of estimations for which the common trend hypothesis is not rejected. Matching on trends for the matching variables also does not change the main conclusions much, except that the later cohorts reduce employment and profits more. We have also extensively discussed SUTVA and can conclude that spillovers play a small role, but that anticipation to treatment might exist.

Our results fit into the literature in two ways. First, the different findings between our matched TWFE method and the more flexible DiD method highlight the importance of the right DiD design and estimator, as heavily discussed in recent literature. The most important difference here is how the counterfactual is composed and constructed. Second, our findings add to the debate on negative and positive side-effects of environmental regulation. Using data up to the end of phase 3 (2020) allows studying heterogeneity over time. We conclude that some worries over employment loss might be warranted by our findings, but they are of temporary nature. Profits seem mostly unaffected, but investments seem to decrease for the most energy-intense firms when regulated.

Research of the EU ETS's effects remains of interest, as longer time series allow for the evaluation of medium and long-term effects. This can be informative to policy makers that consider the implementation or strengthening of environmental policy. Future research will also allow for the analysis of changes in regulatory stringency, which we here already exploited to some extent when discussing SUTVA. Analysis of phase 4 reforms and the high EUA prices as of 2021 might provide new insights.

## References

- Abadie, A. (2005). Semiparametric Difference-in-Differences Estimators. *The Review of Economic Studies*, 72(1), 1–19.
- Abrell, J. (2021). Database for the European Union Transaction Log [Database, available at <https://euets.info>].
- Borghesi, S., Cainelli, G., & Mazzanti, M. (2015). Linking emission trading to environmental innovation: Evidence from the Italian manufacturing industry. *Research Policy*, 44(3), 669–683.
- Calel, R., & Dechezleprêtre, A. (2016). Environmental policy and directed technological change: Evidence from the European carbon market. *Review of Economics and Statistics*, 98(1), 173–191.
- Callaway, B., & Sant’Anna, P. H. (2021). Difference-in-differences with multiple time periods [Themed Issue: Treatment Effect 1]. *Journal of Econometrics*, 225(2), 200–230.
- Colmer, J., Martin, R., Muûls, M., & Wagner, U. J. (2022). *Does pricing carbon mitigate climate change? Firm-level evidence from the European Union Emissions Trading Scheme* (Discussion Paper No. DP16982). CEPR.
- Daw, J. R., & Hatfield, L. A. (2018). Matching and regression to the mean in difference-in-differences analysis. *Health services research*, 53(6), 4138–4156.
- de Chaisemartin, C., & D’Haultfoeuille, X. (2022). *Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey* (Working Paper No. 29691). National Bureau of Economic Research.
- Dechezleprêtre, A., Gennaioli, C., Martin, R., Muûls, M., & Stoerk, T. (2019). *Searching for carbon leaks in multinational companies* (Discussion Paper No. 1601). Centre for Economic Performance, LSE.
- Ellerman, A. D., & Buchner, B. K. (2008). Over-allocation or abatement? A preliminary analysis of the EU ETS based on the 2005–06 emissions data. *Environmental and Resource Economics*, 41(2), 267–287.
- European Commission. (2004). Commission decision of 29 October 2004 concerning the temporary exclusion of certain installations by the Netherlands from the Community emissions trading scheme pursuant to Article 27 of Directive 2003/87/EC of the European Parliament and of the Council [Decision C(2004)4240-3].
- European Commission. (2005). Commission decision of 22 March 2005 concerning the temporary exclusion of certain installations by the Netherlands from the Community emissions trading scheme pursuant to article 27 of directive 2003/87/ec of the European Parliament and of the Council [Decision C(2005) 866 final].
- European Parliament, Council of the European Union. (2003). Establishing a scheme for greenhouse gas emission allowance trading within the community [Directive 2003/87/EC].
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing [Themed Issue: Treatment Effect 1]. *Journal of Econometrics*, 225(2), 254–277.

- Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies*, 64(4), 605–654.
- Hintermann, B. (2010). Allowance price drivers in the first phase of the EU ETS. *Journal of Environmental Economics and Management*, 59(1), 43–56.
- Jaraite-Kažukauske, J., & Di Maria, C. (2016). Did the EU ETS make a difference? an empirical assessment using Lithuanian firm-level data. *The Energy Journal*, 37(1).
- Klemetsen, M., Rosendahl, K. E., & Jakobsen, A. L. (2020). The impacts of the EU ETS on Norwegian plants' environmental and economic performance. *Climate Change Economics*, 11(01).
- Löschel, A., Lutz, B. J., & Managi, S. (2019). The impacts of the EU ETS on efficiency and economic performance – An empirical analyses for German manufacturing firms. *Resource and Energy Economics*, 56, 71–95.
- Marin, G., Marino, M., & Pellegrin, C. (2018). The impact of the European Emission Trading Scheme on multiple measures of economic performance. *Environmental and Resource Economics*, 71(2), 551–582.
- Martin, R., Muûls, M., De Preux, L. B., & Wagner, U. J. (2014). Industry compensation under relocation risk: A firm-level analysis of the EU emissions trading scheme. *American Economic Review*, 104(8), 2482–2508.
- Petrick, S., & Wagner, U. J. (2014). *The impact of carbon trading on industry: Evidence from German manufacturing firms* (Working paper No. 2389800). Available at SSRN.
- Porter, M. E., & Van der Linde, C. (1995). Toward a new conception of the environment-competitiveness relationship. *Journal of Economic Perspectives*, 9(4), 97–118.
- Sant'Anna, P. H., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1), 101–122.
- Teixidó, J., Verde, S. F., & Nicolli, F. (2019). The impact of the EU Emissions Trading System on low-carbon technological change: The empirical evidence. *Ecological Economics*, 164 (Article Number: 106347).
- Verde, S. F. (2020). The impact of the EU Emissions Trading System on competitiveness and carbon leakage: The econometric evidence. *Journal of Economic Surveys*, 34(2), 320–343.

## A Data details

### A.1 Statistics Netherlands (CBS)

The units in the CBS data are partially constructed by CBS itself. Especially the Business Unit (BE) is a construct that is generated by CBS. Here we will discuss how these units are constructed.

#### A.1.1 Business Unit (BE)

The business unit (BE) captures outward-facing (i.e. non-internal) Dutch production or service-provision that can be seen as one unit. This means that legal firm structures are grouped by purpose into BEs, e.g. a unit producing wooden furniture. This provides several advantages and disadvantages. The main advantage is that the BE is a unit structure that captures economic activity well. Legal firm structures often only exist for fiscal reasons and do not represent economic activity or choices well. The disadvantage is that BEs are constructed and that their composition can change over time, even though these changes might be representative of economic activity within the BE.

### A.2 EU ETS

For the data on the EU ETS, coming from [EUETS.INFO](http://EUETS.INFO), a few transformations are needed.

The main problem occurs when installations change owner. This event is poorly captured by the data and therefore requires manual corrections. The corrections of ownership change were done in the following steps.

1. From the European Commission's Union Registry the lists of (stationary) installations for each phase are downloaded.<sup>10</sup>
2. The owners of each installation are compared across phases. If the owners are unchanged between phases, they are assumed to have been the same within that phase.
3. For the installations of which owners have changed between phases, we search the internet for further information to determine whether there was a transfer of ownership and between whom. From sources like news articles or websites that provide ownership data, we deduce when ownership has changed and to who. Two common situations occur, namely (1) ownership of installations is transferred within a firm group, which effectively means the installation

---

<sup>10</sup>These lists can be found for Phase 1, 2 and 3 on the EC's [website](http://EUETS.INFO).



has the same ultimate owner and (2) another firm purchases the installation, sometimes because the previous owner went bankrupt.

4. For installations that saw their owner change but for which we find no information when this took place, we assumed the change to take place on the day the new phase started.

The dates of ownership change then have to be reconciled with the annual data. For this, the year was chosen in which the ownership change has taken place and this year is considered to be the year in which the new owner takes economic responsibility of the installation.

### **A.3 Details on merging the EUTL with CBS data**

Data that is imported into the CBS environment and that is identified on the chamber of commerce (in Dutch: KvK) number, like the ETS data, is encrypted on the same level. So installations under the EU ETS are imported into the CBS environment and encrypted. Encrypted chamber of commerce numbers can then be used to link EU ETS regulation to the business units.

Based on this encryption, one can find the corresponding CBS person (Dutch: persoon) in each year. This CBS person presents a layer in between the detailed KvK number and the final identifier level, business units (BEs). The CBS persoon itself is just a one to one linking from the KvK number to a CBS internal identifier. In some rare years a KvK number is assigned to two CBS persons within a year. This is because CBS draws from multiple sources which can cause duplicate links. In these cases, we have decided to assign the KvK number to the later created CBS person within that year.

The original ETS plant is thus assigned to a BEID in each year, ownership changes between years are thus uncritically represented here. However, in some years a CBS person is assigned to two BEIDs, which can happen if ownership changes within a year. In these cases, we assign the later BEID to the plant.

The CBS data sets are all identified on the BEID level and so we can in the next step merge the ETS plants to the CBS data sets. In each of these steps some of the companies cannot be assigned to another identifier or data set, such that in the end not all ETS firms can be merged. There is, however, no systemic bias in this. After consultation with CBS, the majority of the firms that we were not able to link stem from site that has merged several ETS installations under one account holder, which are then impossible to link to the BEID in our data.

## B Technicalities of estimation strategy

### B.1 Further explanation and definitions of the group-year specific ATT

We here give the definitions of the inverse probability and outcome regression adjustments as well as their underlying interpretation.

$$\hat{w}_{jc}^{treated} = \frac{G_{jc}}{\frac{1}{N} \sum_i G_{jc}} \quad (\text{B.1})$$

$$\hat{w}_{jc}^{control} = C_{jc} \frac{\frac{p_{jc}(X_j, \hat{\pi}_c)}{1-p_{jc}(X_j, \hat{\pi}_c)}}{\frac{1}{N} \sum_i \frac{p_{jc}(X_j, \hat{\pi}_c)}{1-p_{jc}(X_j, \hat{\pi}_c)}} \quad (\text{B.2})$$

with  $G_{jc}$  being a dummy for if a firm is in the respective treatment group or not,  $C_{jc}$ , a dummy that is one if the firm can serve as a control for that treatment cohort, thus incorporating never treated as well as not yet treated firms, and  $p_{jc}$  as the estimated propensity score for each firm (giving the probability of being in that treatment cohort), based on the controls and the estimated coefficients  $\hat{\pi}_c$  from a logistic regression model. This procedure thus weights controls that are more likely to be treated higher than firms that are unlikely to be treated.

$\hat{m}_{jc}(X_j, \hat{\lambda}_{ct})$  is the estimator of  $\mathbb{E}[Y_t - Y_{base} | X, C = 1]$ . It is thus the difference in predicted values between year  $t$  and the base year for the treated firms, if they were untreated. One thus runs  $y_{jt} - y_{jb} = \lambda X_j + \varepsilon_j$  only on the sample of the untreated units, to estimate the change in outcomes that can be predicted by the covariates and then uses this  $\hat{\lambda}$  to predict  $\hat{m}_{jt}(X_j, \hat{\lambda}_{tc}) = \widehat{y_{jt} - y_{jb}}$ , in this case both for the treated and untreated units.

## C Matching

Our matching algorithm for the TWFE estimation is presented in [Algorithm 1](#). The algorithm is designed to match treated firms to similar enough control firms in order to make a sensible comparison between their economic outcomes. It also attempts to filter for good data quality, e.g. by only considering firms that are observed for several consecutive years around treatment.

### Algorithm 1: Matching

1. Enforce common support between treated and control units

- (a) For each baseline year, we drop all observations that are outside the common support of the treated and control group.
2. Select treatment period
    - (a) Take treatment period  $T \in T^p$ , where  $T^p$  is the set of treatment periods, i.e. the years 2005, 2008 and 2013 for phase 1, phase 2 and phase 3 ( $p$ ) in the EU ETS respectively.
  3. Select observations to be potentially matched
    - (a) From the ever-treated EU ETS firms, select only those observations that are first regulated in phase  $p$ . Keep all observations from the never-treated group.
    - (b) Only keep units that are observed for all of the years in  $(T - pre, T + post)$ , where we set  $pre = 2$  and  $post = 3$ . This guarantees that resulting matches can be observed around the treatment period.
    - (c) Select only the observations at  $T - pre$ , dropping the panel structure. This year will be the pre-treatment matching period.
  4. Similarity scoring and match decision
    - (a) Measure the Mahalanobis distance between all observations in the selected sample across treatment status for the variables  $X^m$ .<sup>11</sup>  $X^m$  are the matching variables for which we take the number of employees, turnover, wage expenses, energy expenses, and value added and their squared values. We also restrict matches to be only within a 2-digit sector code. Matches across sectors are not allowed.
    - (b) For each treated unit collect the  $H$  closest neighbors based on the Mahalanobis distance. We opt for  $H = 5$  and we do allow for replacement. We also allow for ties, meaning ties are not randomly broken but rather all are included in the result. For the implementation of this step and the previous step we leverage on the `Matching` package's `Match` function in R.
  5. Store matching outcome
    - (a) Remaining matches are stored under matching year  $T - pre$ .

---

<sup>11</sup>The Mahalanobis distance between treated ( $T$ ) unit A's covariate vector  $x_A$  and control ( $C$ ) unit B's covariate vector  $x_B$  is given by  $d(A, B) = \sqrt{(x_A^T - \mu^T)S^{-1}(x_B^C - \mu^C)}$ , where  $S$  is the variance-covariance matrix between  $x^T$  and  $x^C$  and where the  $\mu$ s are the means of their respective series. Note that this distance measure is like a variance-corrected normalized Euclidean distance.

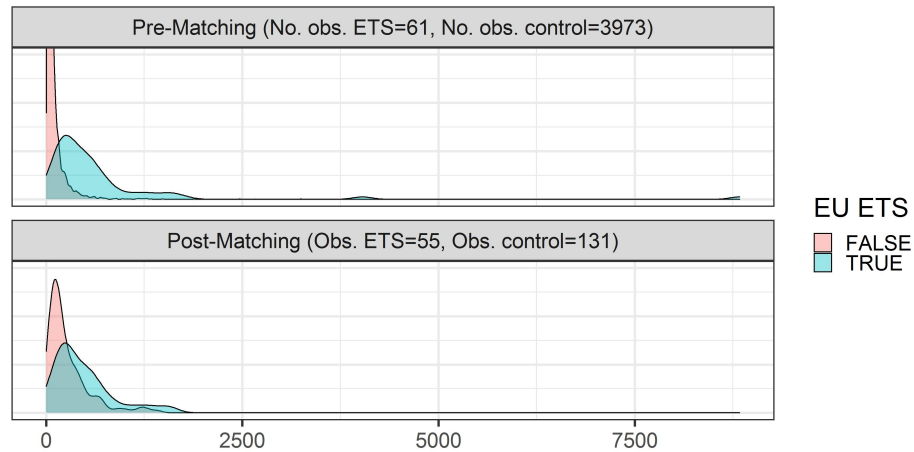
6. Next treatment period

- (a) If not all treatment periods in  $T^p$  are covered yet, select the next value in  $T^p$  and repeat the algorithm from step 2.

Table 1 provides the balancing table after matching. Figure C.1, Figure C.2 and Figure C.3 show the distributions of selected variable for regulated versus non-regulated firms before and after the matching procedure for the pre-phase 1 year 2003, pre-phase 2 year 2006 and the pre-phase 3 year 2011 respectively.

### Distributions for No. of employees

Year = 2003



(a) *Number of employees*

### Distributions for Turnover (Millions EUR)

Year = 2003

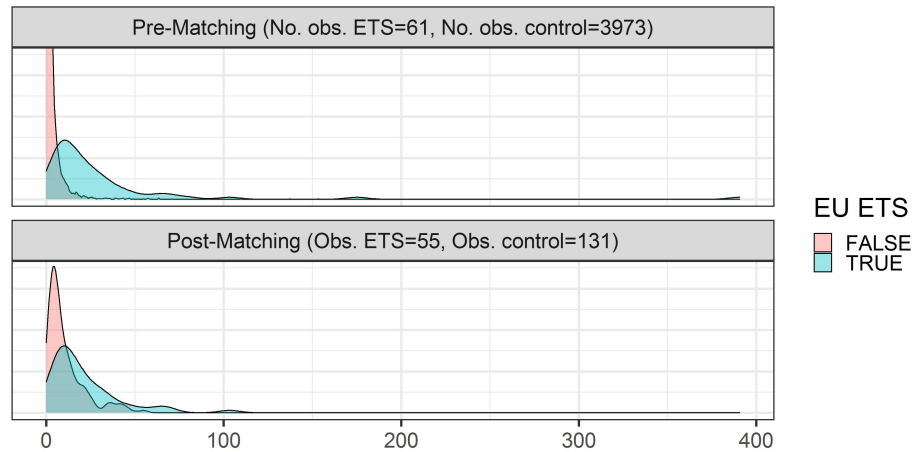


(b) *Turnover*

**Figure C.1:** *Distributions of variables before and after matching for treated and control firms in 2003.*

### Distributions for Wages (Millions EUR)

Year = 2003



(c) *Total wages*

### Distributions for Energy expenses (Millions EUR)

Year = 2003

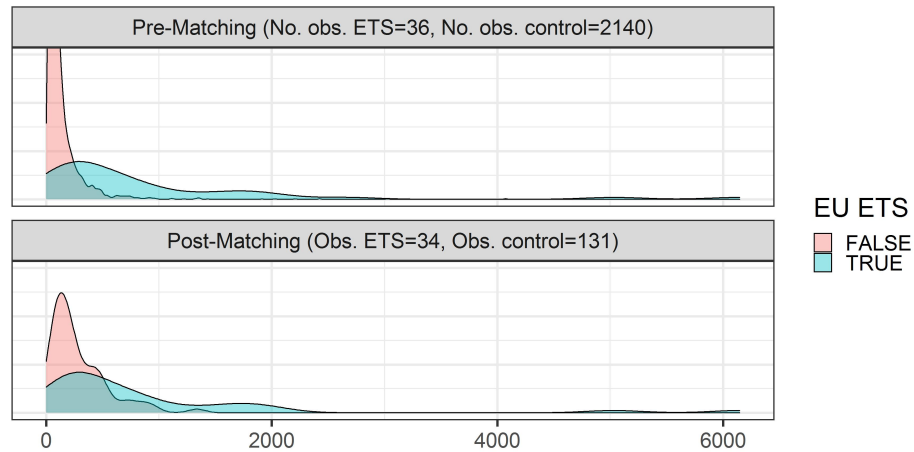


(d) *Energy expenditures*

**Figure C.1:** *Distributions of variables before and after matching for treated and control firms in 2003. (Cont'd.)*

### Distributions for No. of employees

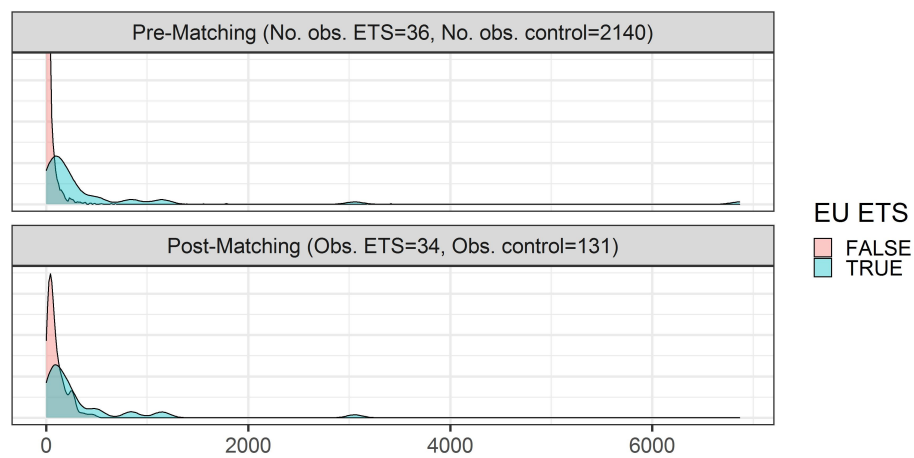
Year = 2006



(a) *Number of employees*

### Distributions for Turnover (Millions EUR)

Year = 2006

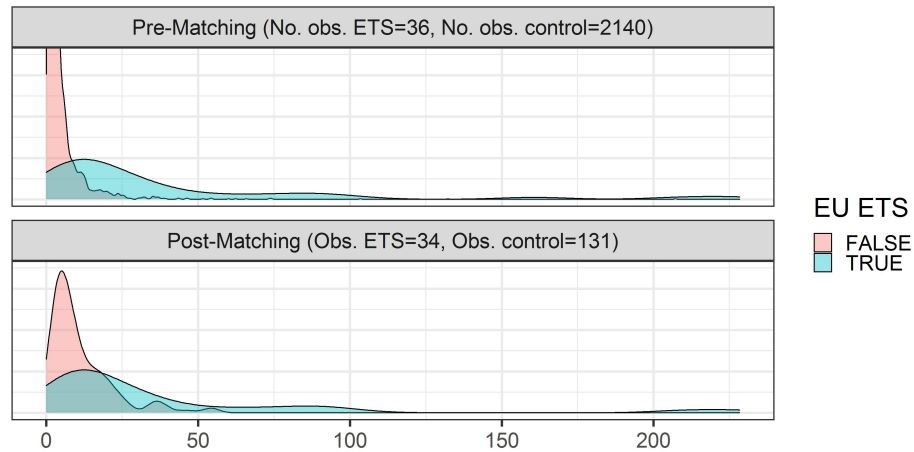


(b) *Turnover*

**Figure C.2:** *Distributions of variables before and after matching for treated and control firms in 2006.*

### Distributions for Wages (Millions EUR)

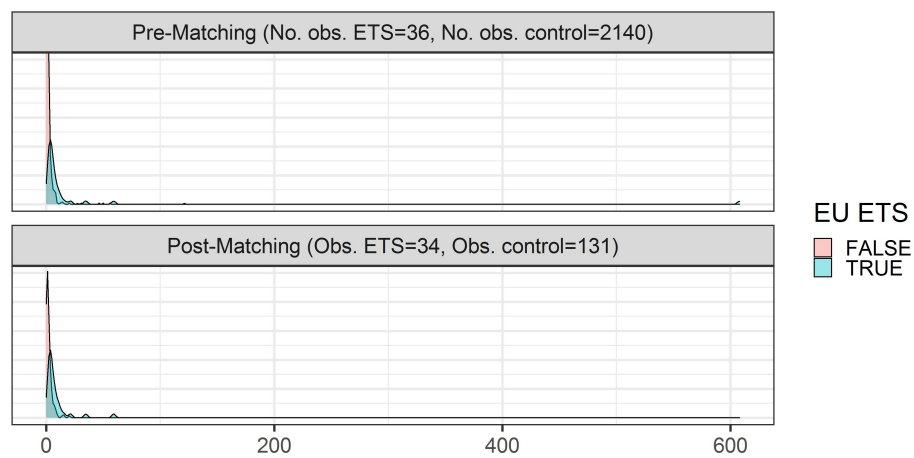
Year = 2006



(c) *Total wages*

### Distributions for Energy expenses (Millions EUR)

Year = 2006



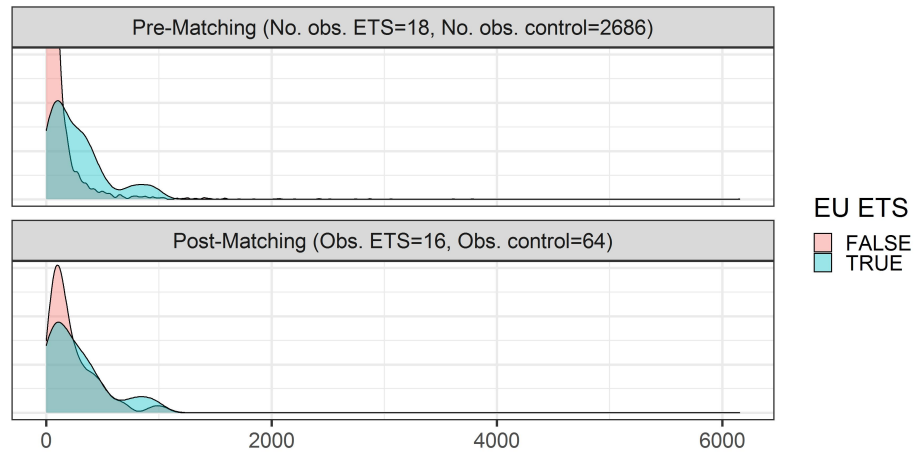
(d) *Energy expenditures*

**Figure C.2:** *Distributions of variables before and after matching for treated and control firms in 2006. (Cont'd.)*



### Distributions for No. of employees

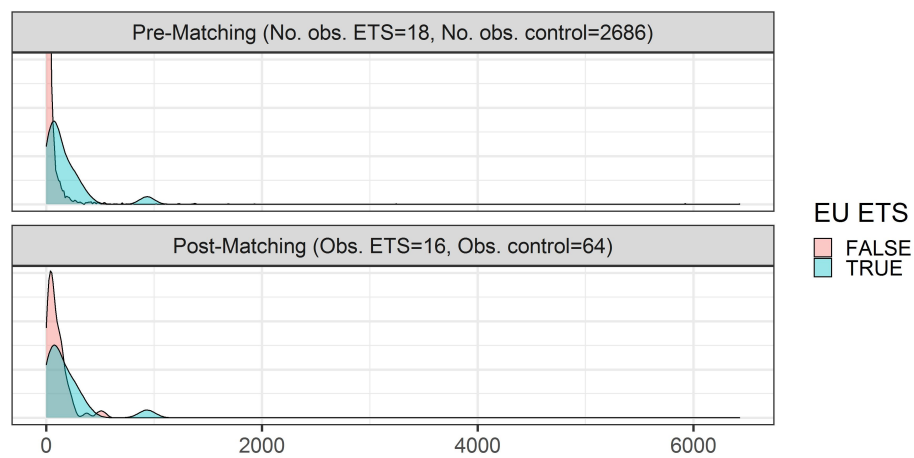
Year = 2011



(a) *Number of employees*

### Distributions for Turnover (Millions EUR)

Year = 2011

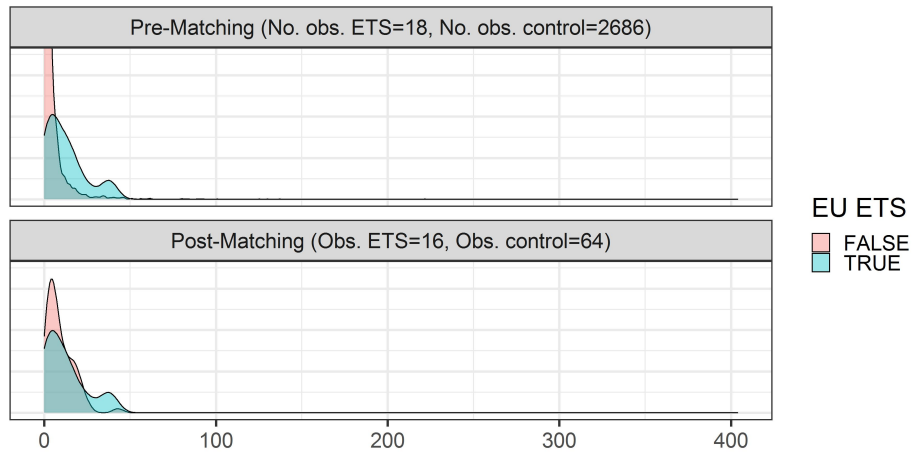


(b) *Turnover*

**Figure C.3:** *Distributions of variables before and after matching for treated and control firms in 2011.*

### Distributions for Wages (Millions EUR)

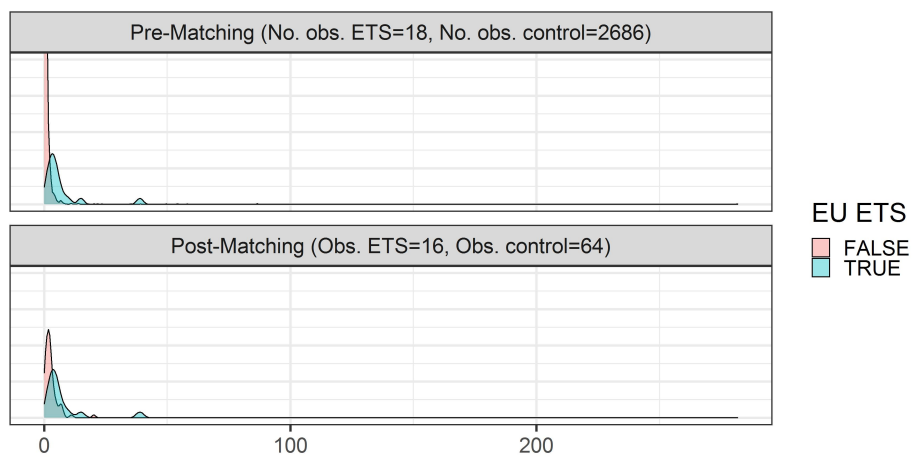
Year = 2011



(c) *Total wages*

### Distributions for Energy expenses (Millions EUR)

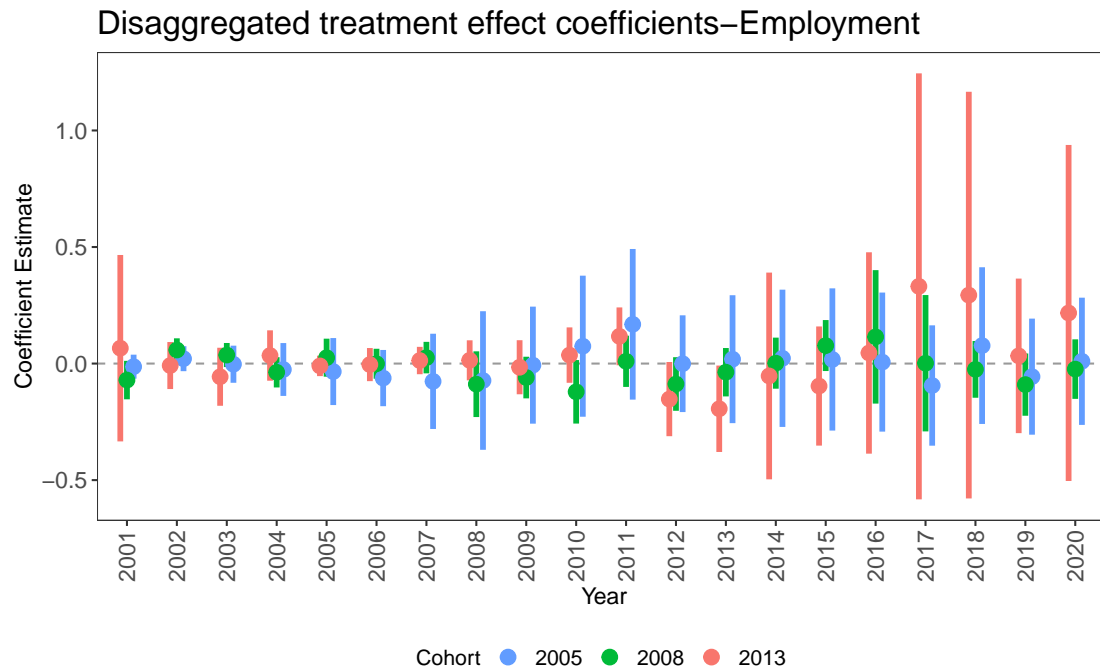
Year = 2011



(d) *Energy expenditures*

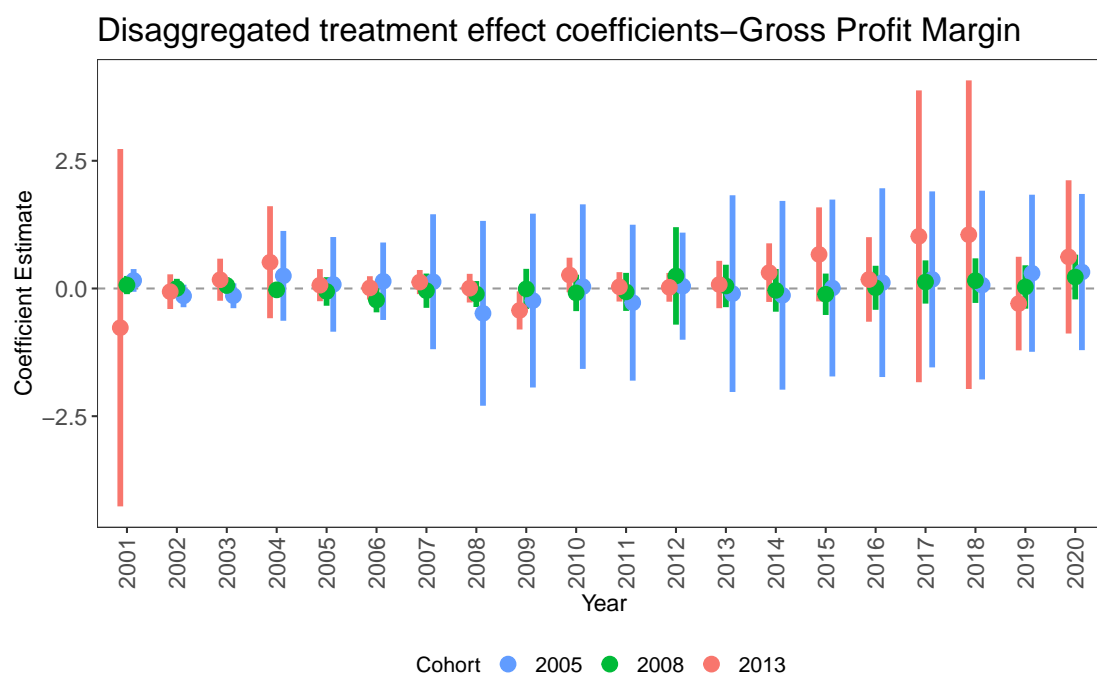
**Figure C.3:** *Distributions of variables before and after matching for treated and control firms in 2011. (Cont'd.)*

## D Additional tables and figures



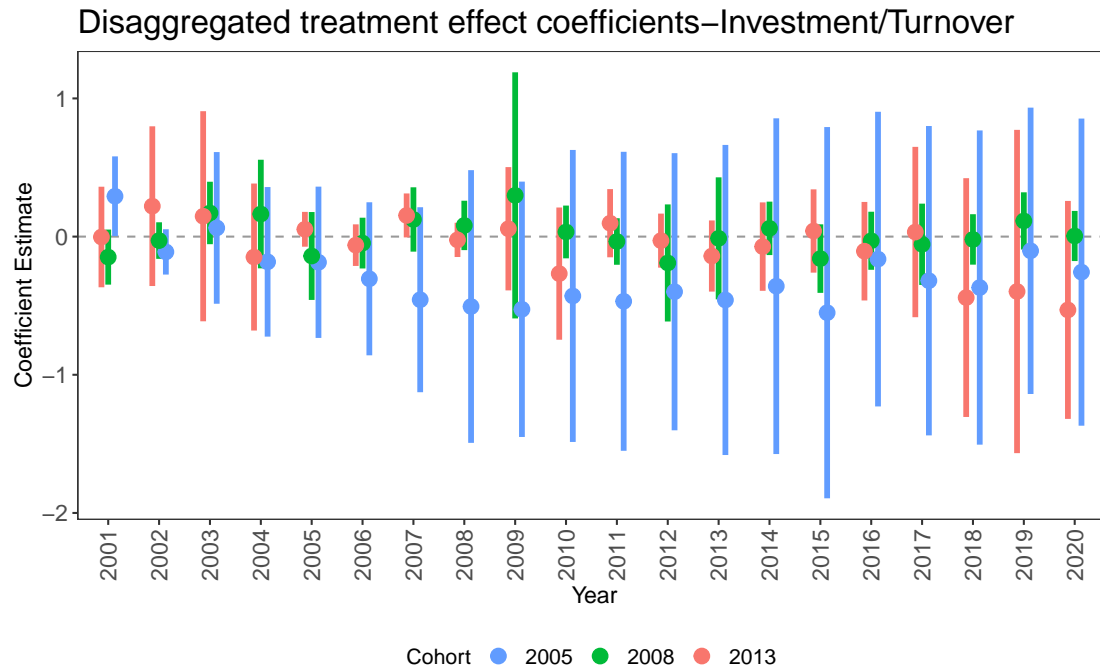
**Figure D.1:** *Treatment effect estimates for employment.*

*Note:* Non-aggregated coefficient estimates for employment from analysis of [Equation 4.3](#). Bars present 95% confidence regions. Pre-treatment estimates are placebo estimates.



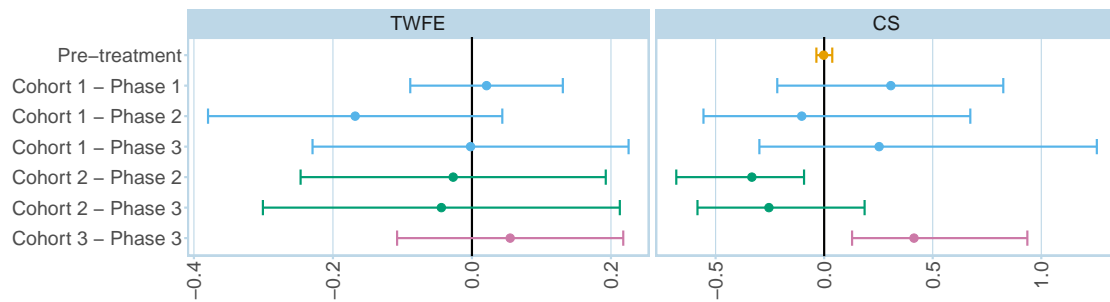
**Figure D.2:** *Treatment effect estimates for employment.*

*Note: Non-aggregated coefficient estimates for gross profit margin from analysis of [Equation 4.3](#). Bars present 95% confidence regions. Pre-treatment estimates are placebo estimates.*



**Figure D.3:** *Treatment effect estimates for employment.*

*Note: Non-aggregated coefficient estimates for turnover from analysis of Equation 4.3. Bars present 95% confidence regions. Pre-treatment estimates are placebo estimates.*



**Figure D.4:** *EU ETS effects on EBITDA margin.*

*Note: Cohort-phase coefficient estimates for cohorts 1-3 for EBITDA margin. Both the TWFE and CS results are provided. Whiskers indicate 95% confidence intervals, which for CS are bootstrapped. Colors refer to the different cohorts.*

# Competitiveness and investments under emissions trading

Leon Bremer<sup>\*‡</sup>      Konstantin Sommer<sup>†‡</sup>

September 2, 2022

## Abstract

We study the effects of the EU Emissions Trading System (ETS) on employment and profits as well as on the investment decisions of Dutch manufacturing firms. Motivated both by sizable differences between firms that are regulated in different phases and by the gradual increase in regulatory stringency, we pay close attention to treatment effect heterogeneity between firms and over time. We use microdata from Statistics Netherlands to apply two difference-in-differences (DiD) estimators: (1) a matched two-way fixed effects regression and (2) a recently developed, more flexible DiD method, designed for staggered treatment and treatment effect heterogeneity. We find that firms that were first regulated in phase 1 and 2 experience temporary employment losses of between 7 to 9% early in the regulation, but we do not find conclusive evidence for changes in profits. Firms that were regulated the earliest reduced their investments throughout all phases.

**JEL codes:** H23, L51, Q52

**Keywords:** Emissions trading, Environmental regulation, Staggered Difference-in-Differences, Treatment heterogeneity, Manufacturing

---

\*Vrije Universiteit Amsterdam, Department of Spatial Economics, De Boelelaan 1105, 1081 HV Amsterdam, and Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam. E-mail: [l.bremer@vu.nl](mailto:l.bremer@vu.nl).

†University of Amsterdam, Department of Macro and International Economics, Roetersstraat 11, 1018 WB Amsterdam, and Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam. E-mail: [k.h.l.sommer2@uva.nl](mailto:k.h.l.sommer2@uva.nl). (corresponding author)

‡The authors are grateful for funding by the A Sustainable Future (ASF) platform from the University of Amsterdam and the travel grant provided by the International Association for Applied Econometrics (IAAE) to its 2022 conference. The authors are also grateful for the opportunities to present at various conferences and seminars, and the feedback received by various participants of these sessions.

# 1 Introduction

This paper studies the effects of the European Union Emissions Trading System (EU ETS) on both the competitiveness as well as the investment decisions of regulated manufacturing firms. We show that firms that became regulated in different years differ from each other, and that regulation stringency significantly varied over time. We carefully disentangle the effects on these different cohorts of firms, and analyse the dynamics of the policy’s effects over time.

The EU ETS is the world’s largest cap-and-trade system, aiming to reduce the EU’s emissions of greenhouse gases. It was implemented in 2005 and caps the amount of emissions within the EU for the covered plants, which together make up for about half of the EU’s emissions. Large emitters have to surrender one allowance for each emitted ton of CO<sub>2</sub>-equivalent at the end of each year. These allowances can be traded on financial markets, thereby establishing a price for carbon. ETS regulation has been amended throughout its four phases (2005-07, 2008-12, 2013-20, 2021-30), in each of which additional installations were regulated and in which the regulation and its stringency were adapted.

Since the beginning of the ETS, policy makers and industry representatives have expressed concerns about the problem of (carbon) leakage. This would occur if EU producers lose competitiveness towards their non-EU competitors due to compliance with the ETS, leading to a loss of jobs in Europe while not reducing global emissions, but merely shifting them elsewhere. This would be undesirable both from an EU welfare, and from an environmental point of view. Meanwhile, stricter regulation could also incentivize firms to invest in new technologies and gain competitiveness in the long run, a hypothesis in line with Porter and Van der Linde (1995). This study will estimate the causal effect of the ETS on both of these potential policy side effects.

While previous studies have estimated the effects of the ETS, few analyzed more than one dimension of the underlying treatment effect heterogeneity. We focus our analysis on these heterogeneities, motivated by the differences between firms that start being regulated in different phases (which we will call “cohorts”) and by the changing nature of the regulation over time. If the ETS’s effect differs between regulated cohorts, neglecting such heterogeneities might lead to biased estimates that could explain the absence of significant findings in the literature.

Furthermore, few studies cover the more recent years in which allowance prices rose and in which amendments like the Market Stability Reserve (MSR) were introduced and implemented. This paper tries to close these gaps.

We are able to use detailed firm-level microdata from Statistics Netherlands (in Dutch: CBS), the Dutch national statistics agency, and link those to the European Union Transaction Log (EUTL) for information on regulated ETS firms. Besides the more classic difference-in-differences (DiD) implementation through matched

two-way fixed effects (TWFE), we employ a recent, more flexible DiD method developed in Callaway and Sant’Anna (2021) (which we will refer to as “CS”) that is designed for settings with staggered treatment timing, like ours. The latter method is motivated by a recent stream of econometric literature revealing several flaws of TWFE estimators in staggered settings with treatment heterogeneity.

In both approaches, we aim to disentangle the average treatment effect on the treated (ATT) for each cohort-phase combination. This offers ways to break down the effects of the EU ETS between its phases, as well as between its participants. As cohorts could differ from each other, it could imply that these firms also exhibit different treatment effects from the regulation. It also seems likely that effects are not constant over time, as firms might adjust to the regulation and because the nature of the ETS has changed over time, especially in between phases.

The preferred CS method shows negative immediate, but temporary, effects from the EU ETS on employment. Both cohort 1 and cohort 2 experience negative employment effects of between 6-7% in their first regulated phase, but they return to normal employment levels afterward.

Investments also seem to suffer under ETS regulation, as cohort 1 firms persistently reduce their investments with 0.32 to 0.47 industry-level standard deviations. For profitability, the results are less conclusive and mostly statistically insignificant. It thus seems that the large firms belonging to cohort 1 and 2 reduce their size without compromising their profit margins.

This implies that, against our intuition, the last phase, which was the most stringent so far, did not lead to larger losses in competitiveness. The fact that we nevertheless establish these negative effects on firms’ investments and (temporarily) on employment could be driven by the export and energy intensity of the Dutch manufacturing sector.

The two methods result in somewhat different findings, as the TWFE regression indicates larger employment losses for the first cohort, but does not indicate significant reductions in investments. These differences are likely to be driven by the fact that the control group in the CS estimation is a better fit, and highlight the importance of the right methodology choice in staggered DiD settings. Testing for the robustness of our results and discussing the underlying assumptions confirms our results.

The paper continues as follows: [Section 2](#) discusses the related literature. The data and policy background are discussed in [section 3](#). The methodology and results are presented in [section 4](#) and [section 5](#) respectively. [Section 6](#) shows the robustness of these results and discusses the assumptions underlying our identification. [Section 7](#) concludes this research.



## 2 Related literature

There are several studies on the effects of the ETS on the competitiveness of regulated firms as well as potentially resulting carbon leakage and related technology adoption. Several of these studies use administrative firm-level data in other countries and apply comparable difference-in-differences methods. Other studies utilize a larger set of EU ETS firms combined with publicly available data sets (e.g. Calel & Dechezleprêtre, 2016). Underlying all studies is the complexity of finding appropriate control firms that are unregulated but sufficiently similar, in order to draw causal conclusions.

Most studies that rely on matching estimate the treatment effect on the treated by using the semi-parametric estimator of Heckman et al. (1997). In this framework Petrick and Wagner (2014) and Jaraite-Kažukauske and Di Maria (2016) find no negative effects of the ETS on productivity and employment for Germany and Lithuania, respectively, and Colmer et al. (2022) find no evidence for outsourcing in France. Löschel et al. (2019) additionally use a two-way fixed effects setting to analyze the ETS's effect on productivity in Germany. The authors interestingly find significant positive effects on productivity using the Heckman-style estimator, but not in the regression estimation. According to the authors this effect is likely driven by a positive EU ETS effect on efficiency in some of the regulated industries. Marin et al. (2018), using non-administrative micro data from Bureau van Dijk for a larger set of countries, also do not find negative effects on economic performance, but do find an increase in labor productivity.

All of these studies, however, only use data for the first phase and some years into the second phase. Since the stringency of the ETS increased significantly in the second and third phases, adding later phases may lead to stronger and more clear-cut results. Two studies that do look at later phases, however, also do not find strong negative effects of ETS regulation. Dechezleprêtre et al. (2019) use data on multinational firms up to 2014 and analyze carbon shifting within these firms, again without finding much evidence of leakage. Klemetsen et al. (2020) do look at phase 3 and analyze firms in Norway. Their regression methodology accounts for different effects between phases, but not between companies starting in different phases. They find a slight increase in productivity in phase 2, but no significant effect in the other phases.

In a literature survey, Verde (2020) comes to the conclusion that there is no convincing evidence of leakage and losses in competitiveness due to the ETS yet. The authors also highlight that this might be due to the short time span covered in almost all studies and point to the importance of analyzing more long-term indicators like investments. This study is trying to address both of these gaps.

When it comes to the ETS's effect on innovation, the literature is smaller, but still contains important contributions. Calel and Dechezleprêtre (2016) show

in a large multi-country panel that the ETS has increased green patenting, and Borghesi et al. (2015) show in an Italian phase 1 firm-level panel that regulated sectors have increased innovation, but that this varied by treatment stringency of the sector. A survey by Teixidó et al. (2019), however, comes to the conclusion that evidence on the ETS’s effect on innovation is still too sparse for a coherent conclusion.

Further, to address time-varying heterogeneity in the treatment effects the econometric design needs to be appropriate. There is recent discussion of DiD and its potentially imperfect approximation using TWFE (see e.g. de Chaisemartin & D’Haultfoeuille, 2022). Callaway and Sant’Anna (2021) offer a more flexible alternative to TWFE. In this paper we will employ their estimator on top of the more classical TWFE estimator.

Our contribution to the previous work is threefold. First, we add analytically to the debate about causal effects of the ETS, by adding insights into the potential treatment heterogeneity of the ETS. For this, we employ both a classical matched-TWFE estimator, but adapt it such that it accommodates cohort-phase specific estimates, as well as a recently-developed more detailed semi-parametric estimator. Second, we benefit from longer time series, allowing us to estimate the later phase’s effects, as well as the longer term effects from earlier phases. Third, we are able to use detailed administrative data. We have access to data on investments, on top of the more traditional indicators for firm performance and competitiveness. Additionally, the Netherlands are due to their export orientation and rather energy-intensive industrial structure a country in which competitiveness effects might be more clearly visible.

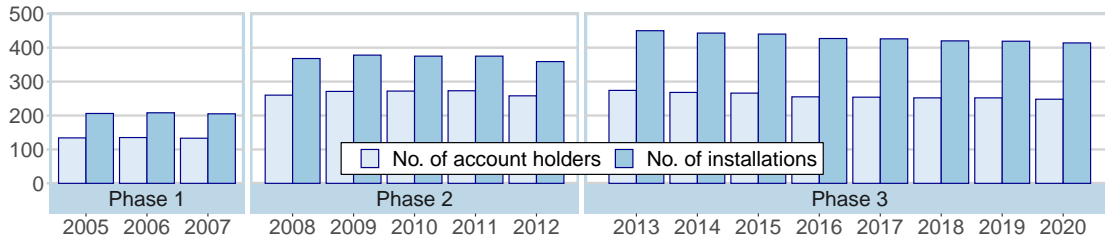
## 3 Data and policy background

### 3.1 EU ETS policy background

The EU ETS regulates installations, which we will also refer to as plants. Each of these plants is registered under one owner, the account holder, at a time in the European Union Transaction Log. The amount of regulated active installations in the Netherlands and their account holders can be found in [Figure 1](#).<sup>1</sup> After its initial implementation in 2005 the ETS has been largely revised 3 times when new phases came into effect, in 2008, 2013 and 2021. Most of these revisions aimed at making the system more restrictive and effective. This study uses data until 2020, thereby excluding Phase 4.

---

<sup>1</sup>Only (former) Operator Holding Accounts that are registered in the Netherlands are selected. The connected installation must have positive verified emissions for that year.



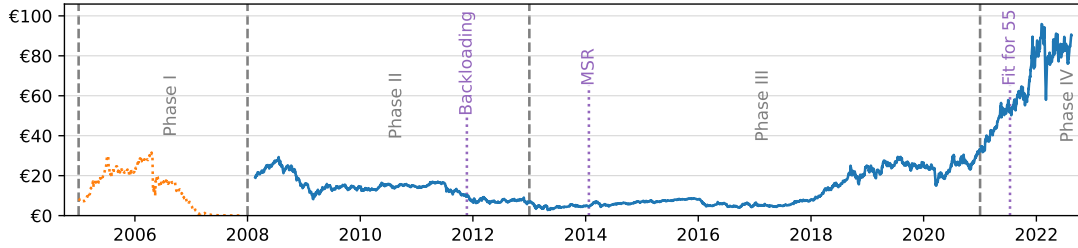
**Figure 1:** *Account holders and installations regulated under the EU ETS.*

*Note:* Number of active installations regulated under the EU ETS in the Netherlands and their account holders. Source: authors calculations based on EUTL data accessed through [EUETS.INFO](http://EUETS.INFO).

In phase 1 (2005-2007) allowances were handed out so plentiful that their price dropped to zero towards the end of the phase, see [Figure 2](#). In the Netherlands actual emissions were almost 15 percent below the number of allocated allowances (Ellerman & Buchner, 2008). Phase 2 (2008-2012) added nitrous oxide as a greenhouse gas and increased the penalty for non-compliance from €40 to €100 per tonne of CO<sub>2</sub>-equivalent. The amount of regulated installations within the Netherlands increased from 205 to 368 (see [Figure 1](#)), mostly because in Phase 1 150 Dutch installations were excluded from the ETS.<sup>2</sup>

Even more greenhouse gases were added in Phase 3 (2013-2020). Also, the default allowance allocation method switched from grandfathering to auctioning. To counteract the low emissions prices that were arguably caused by sustained low demand, the European Commission implemented two sets of new rules to the ETS. First, starting from 2014 the auction of new allowances was postponed until 2019-2020, which was referred to as Backloading. Second, in 2019 the Market Stability Reserve (MSR) started operating. The MSR takes the backloaded allowances and puts them in a reserve. Depending on demand and supply, allowances will be added to the reserve or released from the reserve. As of 2023 excess allowances in the reserve might be permanently cancelled. Further, manufacturing sectors in the aluminium and chemicals production were added to the coverage. This did not change the number of regulated account holders much, but it significantly increased the number of regulated plants (see [Figure 1](#)). Arguably, more plants of the same owners were regulated in phase 3.

<sup>2</sup>The following decisions by the European Commission (EC) provide further details of the phase 1 exemptions for Dutch installations. In October 2004 the EC exempted 93 installations and in March 2005 the EC exempted a further 57 installations (European Commission, 2004, 2005). Other countries that have exempted some firms from the regulation in the first phase were the UK, Sweden and Belgium. More information can be found on the [EU Commission's website](http://EUCommission's website)



**Figure 2:** *EU ETS allowance price.*

*Note: The EU ETS’s allowance price in Euros per tonne of CO<sub>2</sub>-equivalent. These are day closing prices for its futures contracts. The futures montage ECF00-NDEX is plotted in solid blue, and this data is accessed through FactSet. The December 2007 futures price for phase 1 allowances is plotted as a dotted orange line. These allowances were not transferable to later phases. The phase 1 data come from the [European Environment Agency](#). Vertical dashed grey lines indicate the starts of a new phase, while purple dotted vertical lines indicate early proposal dates of amendments to the EU ETS.*

Phase 4 (2021-2030) mainly sped up the rate at which the cap decreases over time, the Linear Reduction Factor (LRF), and it strengthened the MSR.<sup>3</sup>

The changing degree of stringency is also reflected in the allowance price path, as depicted in [Figure 2](#).<sup>4</sup> Prices decreased to zero at the end of phase 1, then started around €20 in phase 2, but stayed around only €10 for several years. Even though economists argue about the optimal price of carbon, such low prices have almost uniformly been deemed as too low to have the intended impact. Prices have started to increase since 2017/2018 and have nearly reached €100 in 2022, making the ETS far more restrictive in recent years.

### 3.2 EUTL and Dutch microdata

The data for this project comes from two main sources. First, the European Union Transactions Log data is accessed through [EUETS.INFO](#), a free service that provides cleaned data from the EUTL (Abrell, 2021). Second, Dutch firm-level data is accessed through the microdata services of Statistics Netherlands (in Dutch: CBS).

The data collected from the EUTL contain information on the free allocations of allowances, verified emissions, allowances surrendered, and the use of interna-

<sup>3</sup>Please refer to the [European Commission’s webpage](#) for more details.

<sup>4</sup>ETS stringency is not the only driver of the allowance price. A body of literature studying the ETS price drivers has identified fossil fuel prices to play a key role (see e.g. Hintermann, 2010).

tional credits, both by installation and account holder. Account holders in the EUTL can potentially own several regulated plants and are registered under a national identification number.<sup>5</sup> The data are organized in an unbalanced panel spanning the years 2005-2020 and a total of 439 unique account holders, owning 598 installations.

The CBS data are not publicly accessible and are anonymized. They contain rich firm-level information on economic activity of almost the entire population of Dutch firms with more than 50 employees. The data contain information like the number of employees, costs of goods sold and turnover, as well as investment data. This study is restricted in scope to manufacturing firms and relies on more than 40,000 firms over a time span of 21 years. To deflate monetary variables, we use Eurostat’s industry producer price index for the Netherlands.

We link EUTL data to the administrative firm-level data of CBS. The linking takes place by the use of the chamber of commerce identifiers that are available in the EUTL for Dutch account holders and in the CBS data. Within CBS, several chamber of commerce numbers can comprise a “business unit”, a construct defined by CBS and further explained in [Appendix A.3](#). We will from here on refer to these business units as “firms”. After linking the EUTL data to CBS’s anonymized data we are not able to identify individual firms anymore.

As a business unit can comprise multiple account holders and plants, it can be the case that a business unit is regulated through more than one plant. We do not make a distinction here and consider each business unit (firm) as regulated if it owns at least one regulated plant in that year. Our level of analysis is on this business unit level, referred to as the firm level.

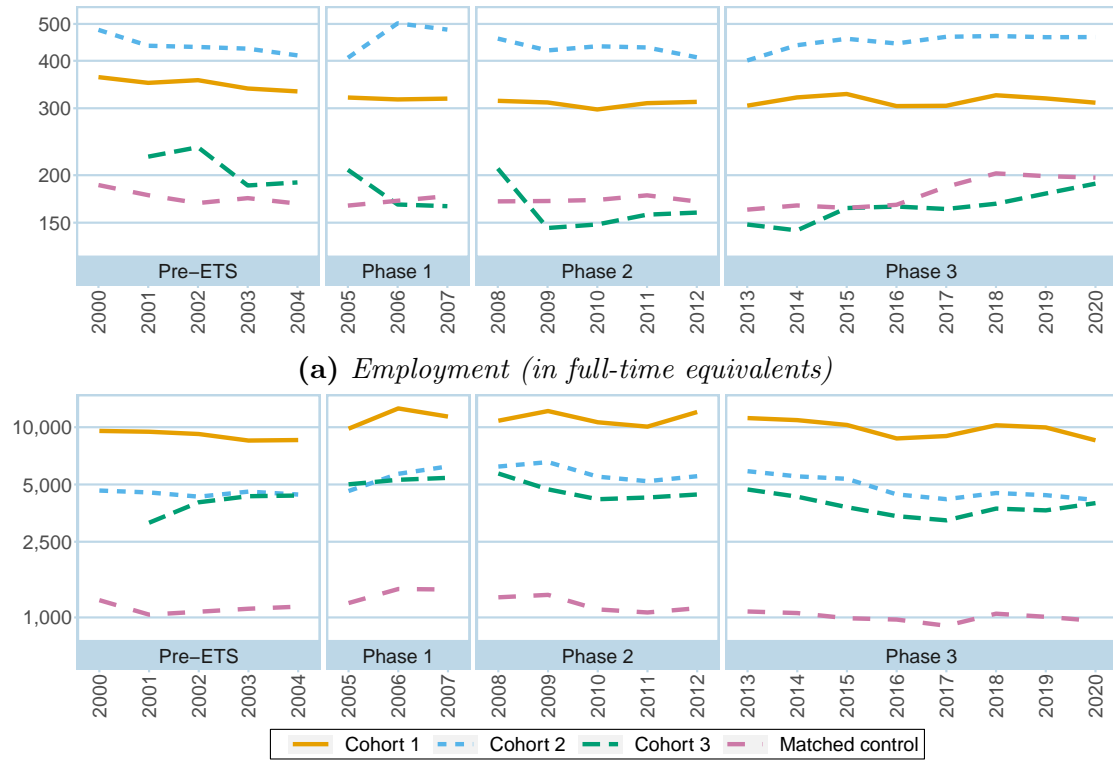
### 3.3 Descriptives

In this section we will elaborate on the two most important sources of heterogeneity that this study tries to disentangle. First, we show the substantial differences between the regulated cohorts, and then we show the development of the ETS treatment stringency over time.

[Figure 3](#) shows the development of the average firm over time for energy expenditure and employment. The plot shows averages for the different ETS cohorts as well as for a set of matched control firms, that is chosen to be as similar to the treated firms as possible, as is further outlined in [Appendix C](#).

---

<sup>5</sup>As installations are assets, they can be purchased from or transferred to other firms. Such changes of ownership are not perfectly captured by the data. Many installations do not change ownership between EU ETS phases in our data, but for the ones where it does change, we manually look up the date of ownership change using online public sources. Sources can be online news articles or websites that provide information about ownership structures. The list of manually assigned ownership changes and their respective source is available upon request.



(a) *Employment (in full-time equivalents)*

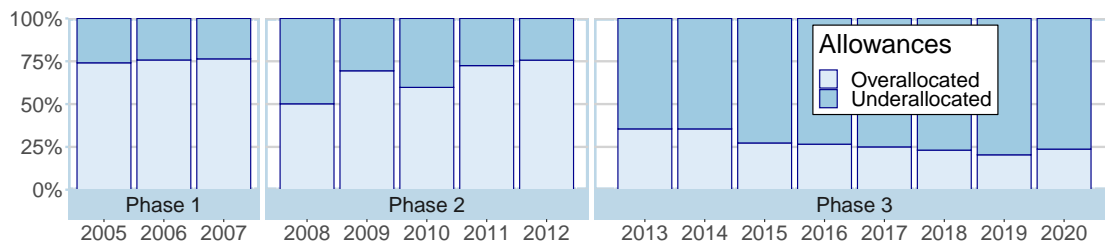
(b) *Energy expenses (in thousands 2015 Euros)*

**Figure 3:** *Averages of selected variables over time.*

*Note: Cohorts 1-3 consist of firms first regulated in Phases 1-3, respectively. The Matched control group consists of unregulated firms that are matched to regulated firms. The vertical axes are on a log scale.*

One can see that firms regulated in 2005 are by far the largest energy consumers. This makes sense, since large emitters were regulated first. Note that these are energy expenses and that energy prices are responsible for some of the time variation. In terms of employment, treated firms seem to be more similar across phases, with now the second cohort being the largest. One can also see that regulated firms are far more energy-intense than unregulated firms, but are more comparable in terms of employment, even though there still remains a gap. It is difficult to hypothesize on the estimated treatment effect, as panel (a) does not show clear kinks at the treatment dates.

Another form of heterogeneity lies in the treatment stringency that a firm experiences from the regulation at different moments in time. As almost all allowances were handed out for free in the first two phases, one could argue that regulation was not stringent in these phases. It was also not uncommon that firms were over-allocated with free allowances, causing these firms to be effectively net subsidized by the ETS. In theory the allocation of allowances should not influence the decision of the firm, as the allocation does not influence the firm's opportunity costs. However, these allocations are likely to have mattered in practice, as firms might have not regarded the regulation as binding and might subsequently not have adjusted their behavior. Profits and financial constraints are mostly unaffected when all emissions are covered by freely allocated allowances.



**Figure 4:** *EU ETS stringency.*

*Note:* The share of Dutch regulated firms that receive more (or less) allowances for free than their verified emissions. *Source:* authors calculations based on EUTL data accessed through [EUETS.INFO](http://EUETS.INFO).

In phase 3, however, allocation mechanisms changed to auctioning as the default option. As many firms were exempted from the switch to auctioning, treatment stringency became more heterogeneous between firms, with some still receiving more allowances than needed, but most now receiving fewer than needed. [Figure 4](#) presents an overview of ETS stringency over time. Together with the changes in prices, shown in [Figure 2](#), this creates significant heterogeneity in the policy stringency both over time and across firms.

## 4 Methodology

### 4.1 Measures of competitiveness and investment intensity

We are interested in the ETS's effects on (1) competitiveness and (2) investments. We measure these concepts with three dependent variables, namely, (1a) the firm's employment and (1b) its profit margins, to inform us on competitiveness, and (2) its investment intensity, to inform us on investment outcomes. Tracking employment outcomes also allows us to evaluate whether domestic environmental regulation indeed led to job losses at home, an often heard counterargument to unilateral environmental policy. Profit margins directly evaluate the profitability of the regulated firms. They also show to what extent regulated firms were able to charge a price that was above their marginal costs, thus they also show if regulated firms were able to pass on additional costs of the regulation to their consumers. This ability likely decreases with the level of competition from abroad. Investment intensity estimates in how far firms are incentivised to invest into new technologies as a response to the regulation.

We measure employment in full-time equivalents (FTE) and use two definitions for the profit margin, (1) the gross profit margin and (2) the EBITDA margin. Gross profits measure the difference between turnover and the costs of goods sold. EBITDA measures earnings before interest, taxes, depreciation and amortization. We scale both measures by turnover to transform them into a margin. Gross profits are generally larger than EBITDA. They are also not influenced by a firm's financial operations, and thereby for example exclude a firm's income from holding activities.

We use investments into all fixed assets and scale them by turnover in order to scale by the size of the firm's activities. As an alternative specification we also scale investments by employment.

In order to further improve comparability across industries, we normalize the profit margins and investment ratios. We do this according to the following normalization:

$$\bar{x}_{jit} = \frac{x_{jit} - \mu_{it}^x}{\sigma_{it}^x}, \quad (4.1)$$

where  $j$ ,  $i$ ,  $t$  refer to firm, industry and time, and  $\mu^x$  and  $\sigma^x$  are  $x$ 's industry-time mean and standard deviation. Note the different subscripts, indicating that the variable is normalized using the industry-specific mean and standard deviation. This way the resulting transformed variable is comparable across industries.

If variables are in monetary terms, they are deflated such that they can be compared over time. For this deflation we use Eurostat's industry producer price index for the Netherlands.



## 4.2 General identification strategy

To identify the effects of the ETS, we use the fact that not all manufacturing firms in the Netherlands are regulated under the ETS. Regulation is on the plant level and there are mainly two criteria for inclusion in the ETS, either (1) through exceeding a certain sector-specific threshold related to energy input or production capacity, or (2) through incorporating specific processes that imply automatic regulation.<sup>6</sup> This implies that one can attempt to find comparable control firms for each treated firm that are both active in comparable production processes and are comparable in terms of size, employment characteristics and energy input.

To then evaluate the causal effects of the EU ETS, we apply two empirical policy evaluation methods. The key in these methods is to use detailed microdata on observed firms to compare the outcomes of treated firms, i.e. firms receiving regulation, to the outcomes of *comparable* control firms, i.e. firms not regulated (yet). The first method is a matched two-way fixed effects regression and the second method is a less parametric DiD design suggested by Callaway and Sant’Anna (2021).

In general two main steps can be identified in this evaluation process, namely (1) matching or weighting, in which we score firms across treatment status based on their similarity, and (2) comparison, in which we either regress our outcome variable on treatment status or take differences in outcome variables across treatment status. The second step utilizes the weights established in the first step. In both methods, we estimate a treatment effect for each cohort-phase combination, thus controlling for the heterogeneity between treated cohorts and in different phases.

## 4.3 Sample selection

One estimation issue arises from firm exit and, to a lesser degree, firm entry, from and to the sample. As we are dealing with anonymized microdata it is not possible to determine if such an exit is due to closure of the firm, an acquisition by another firm or due to changes in the firm structure. To minimize the effect that sample composition could have on our results, we curtail our sample to firms that we observe continuously from two years before to three years after treatment start. Unregulated firms also face this requirement when considered for matching.

We also enforce a common support for all of our covariates (employment, energy costs, turnover and total wage bill) between treated and control groups in the baseline years. The baseline year is two years before treatment start, to allow for

---

<sup>6</sup>For a detailed overview see Annex I of European Parliament, Council of the European Union (2003).

one year of treatment anticipation.<sup>7</sup>

## 4.4 Matched TWFE method

We break the matching and regression up in the following two subsections. The first one explains the matching that provides the weights, and the second one presents the details of the regression.

### 4.4.1 Matching

The goal of matching is to select similar observations across treatment status from the data. In general a matching algorithm provides a similarity score between each pair of observations in the sample data. If provided with  $n$  observations, the matching outcome matrix  $M$  has dimensions  $n \times n$ . For our TWFE application the pair information is dropped and only those observations with a high enough similarity score to any other observation across treatment status are kept, collapsing the matching information from  $M$  to a binary  $n \times 1$  vector, indicating for each firm if it will be kept in the estimation or not. Observations in the non-treated group that do not have a high enough similarity score with a treated observation are thus dropped from our sample. This way matching boils down to sample selection.

The matching outcomes are used to select the sample for our TWFE regression. All observations are kept of firms that are matched, either in the treatment group or the control group (i.e. have a value of 1 in the  $n \times 1$  vector). This effectively is a special form of weighting, as the weights are either 1 (for the matched) or 0 (for the non-matched).

We base our matching on all variables that affect the probability of treatment and try to align it with other studies. We only match within the two-digit industry code, and base the similarity on a firm's employment, energy costs, turnover and total wage bill as well as the squared values of these variables. Matching happens two years before treatment start, to account for anticipation. Our matching algorithm is further elaborated in [Algorithm 1](#) in [Appendix C](#).

---

<sup>7</sup>Before phase 1 the important directive for the establishment of an ETS was passed in 2003, before phase 2 the national allocation plans had to be published in 2006, and before phase 3 the commission passed directive 2011/540/EU in 2011, extending the scope of regulated greenhouse gases and industries.

#### 4.4.2 TWFE regression

Using the resulting matched sample, we can estimate the impact of the EU ETS's phases on each cohort's outcomes. Our two-way fixed effects regression looks as

$$y_{jt} = \sum_{c \in C} \sum_{p \in P} ETS_j^c \times P_t^p \times \mathbb{1}_{\{p \geq c\}} \alpha^{cp} + \gamma_j + \gamma_t + \varepsilon_{jt} \quad (4.2)$$

where  $y$  is the outcome variable of interest and subscripts  $j, t$  refer to the firm and year.  $ETS^c$  is a dummy variable that is equal to one if firm  $j$  is in cohort  $c$ .  $P$  is a dummy that is equal to one if year  $t$  is in ETS Phase  $p$ . As there are three phases in our data range, we have  $C, P \in \{1, 2, 3\}$ . The interactions of the two variables thus present the treatment indicators of our DiD regression. The coefficients of interest are the corresponding  $\alpha$ s, with one coefficient for each of the six post-treatment cohort-phase combinations (i.e. cohort 1-phase 1 through cohort 3-phase 3).

We include firm and optionally year fixed effects, but abstain from including time varying controls, as these are likely to be affected by the treatment itself.  $\varepsilon$  is the error term, which we allow to be heteroskedastic and serially correlated. We estimate the model using ordinary least squares.

The dependent variable  $y$  is either the log-transformed number of employees or a normalized EBITDA margin when interested in the competitiveness effects from the EU ETS, and investments in fixed assets, scaled by turnover and normalized, when interested in the investments response.

### 4.5 CS2021's DiD estimator

Our second approach also relies on matching and a DiD design, but it does so in a less parametric fashion. It follows the approach by Callaway and Sant'Anna (2021). We briefly motivate the use of another estimator by discussing the drawbacks of TWFE, before discussing the estimator itself.

#### 4.5.1 Potential problems with TWFE in staggered DiD

Recent econometric literature has pointed out several problems with TWFE estimations in DiD settings like ours. This literature focuses on the potential biases in TWFE estimators applied to settings with staggered treatment adaption and potentially heterogeneous treatment effects (see e.g. Daw & Hatfield, 2018; de Chaisemartin & D'Haultfoeuille, 2022; Goodman-Bacon, 2021). This is exactly the case in our setting, in which firms get treated in different phases and in which we both expect these different cohorts to react differently and in which we assume the effect to be time (or phase) dependent.

The key problem of TWFE in such cases is that the derived estimator for the Average Treatment Effect on the Treated is a weighted average over the ATTs of the different treatment groups at different times, without explicitly appreciating this and without being able to control the weights of the individual ATTs. The estimator thus only gives a clearly interpretable ATT if treatment effects are constant both over time and between treatment groups.

As described in detail by Goodman-Bacon (2021), these weights can in some cases be negative, resulting in a distorted estimate. Even if the weights are non-negative, they are unobserved and mostly determined by the group size underlying the estimate and the distance to the start of the treatment.

Another problem of the matched-TWFE estimator is that most of the matching information is lost in the regression step. Matching is purely used for sample selection, while the link between matched treated and non-treated units is not taken into account in the estimation. This means that a control firm that is matched to a treated firm in cohort 1 will serve as a control also for treated firms in cohorts 2 and 3, and so on.

#### 4.5.2 The estimator

To address the above mentioned issues, we make use of the estimator developed in Callaway and Sant’Anna (2021). Its main advantage lies in the fact that it estimates ATTs for each treatment cohort – the group of firms starting treatment in the same phase – and at each year into the treatment. It also allows for different aggregations of those estimates, enabling us to restrict the type of heterogeneity.

The estimator is in essence an application of the doubly-robust DiD estimator of Sant’Anna and Zhao (2020) to staggered settings. It pays close attention to the conditioning on covariates, combining both inverse probability weighting (see Abadie, 2005) as well as outcome regression adjustment (see Heckman et al., 1997). The latter is also frequently used in adjusted versions in comparable ETS papers like Martin et al. (2014) or Löschel et al. (2019).

While the inverse probability weighting tries to re-balance the control group based on their probability of being treated, thus in fact on their similarity to the treatment group, the outcome regression adjustment tries to take out trends in the outcome variable that are covariate dependent. The CS estimator is therefore consistent as long as the covariate conditioning is correctly specified by either one (or both) of the two covariate conditioning strategies (therefore referred to as being “doubly-robust”).

An additional advantage of this weighting is that each treated firm is linked to a specific set of control firms and is only compared to these control firms. This is in contrast with the matched TWFE estimator.

The estimator for each cohort,  $c$ , and year,  $t$ , is then a common average treat-

ment effect DiD estimator. It compares the outcome of each firm in year  $t$  to the firm’s own outcome in the base year,  $b$ , and to that of the weighted average difference in outcomes between  $t$  and  $b$  of the respective control group for this firm. For this weighting, both inverse probability weighting and the outcome regression adjustment are used.

The following equation specifies the estimated ATT for cohort  $c$  and year  $t$

$$\hat{\alpha}_{ct} = \frac{1}{N} \sum_{j \in \mathcal{J}} \left[ \underbrace{(\hat{w}_{jc}^{treated} - \hat{w}_{jc}^{control})}_{\text{Inv. prob. weight.}} (y_{jt} - y_{jb} - \underbrace{\hat{m}_{jct}(X_j, \hat{\lambda}_{ct})}_{\text{Outcome reg.}}) \right], \quad (4.3)$$

with  $N$  the number of firms and  $\mathcal{J}$  the set of all firms,  $y_{jt}$  the dependent variable,  $X$  as pre-treatment controls, and  $j, c, t$  referring to firm, cohort and year. *treated* and *control* refer to the treatment status, i.e. regulated or control firms.

$\hat{w}_{jg}^{treated}$  and  $\hat{w}_{jg}^{control}$  are the weights that adjust for the probability of being treated. They are 0 if a firm is not in the respective group and give higher weights to control firms that are more similar to the treated firm, given a set of covariates.  $\hat{m}_{jct}(X, \hat{\lambda}_{ct})$  represents the bias adjustment from an outcome regression, thus deducting the predicted development of  $y$  based on  $X$ , under the assumption that the firm had not been treated. More information on both adjustments and their exact definition can be found in [Appendix B](#).

In this setting, we cannot enforce matching within an industry and thus include industry dummies for the three sectors containing at least one firm regulated in all phases in  $X$ .<sup>8</sup> Employment, turnover, wages and energy expenses as well as their squared values enter as predictors for inverse probability weights and the outcome regression as well. As in the TWFE setting, we assume one year of anticipation, pinning down the base year at two years before the treatment starts. The corresponding standard errors are bootstrapped and clustered at the firm level.

Two sets of candidate control firms can be considered, namely (1) the entire population of firms that has not been treated up to  $t$ , or (2) only the set of firms that will never be treated. For our main specification, we choose to use all not-yet-treated firms as controls, since these will likely be more similar to earlier treated firms. Results for the never-treated control group are presented in the discussion section.

There is no guarantee that this set of control firms exhibits parallel trends in absence of the ETS. To test for parallel pre-treatment trends, we employ a placebo test. By testing whether pre-treatment ATTs (always assuming that the base year is one year before  $t$ ) are different from zero, the test indicates whether

---

<sup>8</sup>These are manufacture of food products and beverages, manufacture of chemicals and chemical products, and manufacture of other non-metallic mineral products. Adjusting the choice of these sectors barely affects our results.

any disparities between the (weighted) treated and control units occurs during the pre-treatment years. We once use a Wald test to for joint significance of these placebo tests and once apply the aggregation outlined in the next paragraph and compute a respective confidence interval to check if it contains zero or not.

Equation 4.3 presents ATT estimations for each cohort-year pair. This allows for more heterogeneity than we assume to be present, as we are interested in cohort-phase effects. Following Callaway and Sant’Anna (2021), we aggregate the separate cohort-year ATTs to a respective cohort-phase,  $\tilde{c}$  and  $\tilde{p}$ , aggregate using cohort-year weights,  $v^{\tilde{c},\tilde{p}}(c,t)$ , defined as:

$$v^{\tilde{c},\tilde{p}}(c,t) = P[t|c = \tilde{c} \text{ and } t \in \tilde{p}] \mathbb{1}_{\{c=\tilde{c}\}} \mathbb{1}_{\{t \in \tilde{p}\}}. \quad (4.4)$$

All weights are non-negative and add up to one within each cohort-phase. These weights are then used in the aggregation

$$\hat{\theta}^{\tilde{c},\tilde{p}} = \sum_{c \in \{1,2,3\}} \sum_{t=2005}^{2020} \hat{v}^{\tilde{c},\tilde{p}}(c,t) \hat{\alpha}_{ct}, \quad (4.5)$$

in which  $\theta$  is the cohort-phase aggregated ATT.

For inference, a bootstrap algorithm calculates a 95% confidence interval around each estimator  $\hat{\theta}^{\tilde{c},\tilde{p}}$ . The algorithm repeatedly draws a subsample from the original sample, keeping the sizes of the cohorts proportional, and estimates the  $\hat{\alpha}_{c,t}$ s and the  $\hat{\theta}^{\tilde{c},\tilde{p}}$ s. From this distribution of estimators the 2.5th and 97.5th percentile determine the 95% confidence interval.

## 5 Findings

This section presents the findings of the proposed methodologies from the previous section. First we discuss the TWFE results together with the associated matching outcomes. We use this opportunity to also discuss the differences between treated and control firms, as the control group is clearer defined in the TWFE than in the CS setting. We then present the results from the CS estimation and subsequently discuss some differences in the results.

### 5.1 TWFE

#### 5.1.1 Matching outcomes

As described in the matching algorithm, we require firms to be observed around the treatment start. We only keep firms that are observed all years from 2 years

before to 3 years after treatment start. This leaves 115 ETS firms in our sample. Of these, 105 are matched to control firms.<sup>9</sup>

Table 1 presents a balancing table, showing the means of the matched firms for the three cohorts. The balancing table shows that regulated firms are larger and invest (relatively) more than unregulated firms. Profit margins are rather comparable between treated and control firms. Even though the matching selects similar firms, some of the size difference remains. This is inherent as the ETS by design regulates larger and more energy-intensive firms. Remaining level differences therefore need to be accounted for within the estimation. The control group of firms in the CS setting will be different and varies each year, as the weights are also based on the outcome regression adjustment. They will, however, be tailored even more to the individual firms thus reducing the size differences.

Further matching outcomes are presented in Appendix C. There Figure C.1 through C.3 provide distribution plots for the matching variables both before and after matching. This is mostly because firms that are very dissimilar to regulated firms are excluded. These figures show that matching improves the comparability of the distributions across treatment status.

### 5.1.2 Regression results

Table 2 presents the results of our TWFE estimation from Equation 4.2 for employment, the gross profit margin and the investment ratio as dependent variables. The regressions include firm fixed effects and either phase or year fixed effects. Cohort and industry instead of firm fixed effects do not lead to substantially different results.

Most of the estimated coefficients are statistically insignificant, indicating that we do not establish many side-effects from ETS regulation in this estimation. The significant estimates, however, have the predicted sign and indicate a reduction in employment for cohort 1 firms. The effect seems to be persistent but is most clearly estimated right at the beginning of the treatment start. On the one hand these firms are the most energy-intensive and are thus most likely to be affected by the regulation, making it reasonable that these firms had to reduce their employment the most as a response to the regulation. On the other hand, the early EU ETS phases are widely deemed to lack stringency, making the result surprising. Cohort 2 shows a similar effect in phase 2, but the estimate is not statistically significant.

The estimated employment effects are also quite sizable, as cohort 1 experiences a nearly 8% decrease in phase 1 compared to pre-treatment years. This effect is mostly maintained in phase 2, as phase 2 employment is 9% lower compared

---

<sup>9</sup>Doing the estimation on a fully balanced panel greatly reduces the power of the estimation, but the qualitative results remain similar.

**Table 1:** *Balancing table for the variables used in the analysis. The cohorts refer to the EU ETS firms' first regulated phase. Treated and control refer to matched EU ETS and non-ETS firms. The groups presented here are the ones in the TWFE estimation. The treated group is almost identical in the CS estimation. Values correspond to the group means two years before treatment start. Monetary variables are in millions of 2015 Euros (Investment/Employment 2015 TEuros). Standard deviations are in brackets.*

	Cohort 1			Cohort 2			Cohort 3		
	Control	Treated	Difference	Control	Treated	Difference	Control	Treated	Difference
Employment	253,78 (242,27)	469,61 (372,46)	215,83 (52,87)	246,86 (226,40)	928,62 (1311,72)	681,76 (225,45)	244,43 (211,32)	276,20 (269,27)	31,76 (68,79)
Gross Profit Margin	0,47 (0,16)	0,51 (0,19)	0,04 (0,03)	0,43 (0,18)	0,57 (0,16)	0,13 (0,03)	0,44 (0,19)	0,45 (0,26)	0,02 (0,07)
Investment/Turnover	0,04 (0,08)	0,06 (0,10)	0,02 (0,01)	0,05 (0,10)	0,05 (0,05)	0,00 (0,01)	0,03 (0,09)	0,06 (0,08)	0,03 (0,02)
Energy costs	2,01 (5,14)	11,98 (14,19)	9,97 (1,95)	2,87 (6,36)	8,24 (11,13)	5,37 (1,95)	2,25 (3,50)	7,09 (9,15)	4,84 (2,30)
Wages	9,53 (9,51)	21,09 (19,71)	11,56 (2,74)	9,49 (8,96)	38,32 (52,65)	28,83 (9,05)	10,80 (10,62)	12,39 (12,16)	1,59 (3,12)
Turnover	84,85 (94,17)	236,19 (329,58)	151,34 (44,90)	97,02 (97,01)	346,64 (566,72)	249,63 (97,40)	117,37 (159,51)	183,15 (226,12)	65,78 (57,53)
EBITDA Margin	0,10 (0,09)	0,12 (0,10)	0,02 (0,01)	0,09 (0,11)	0,12 (0,12)	0,04 (0,02)	0,08 (0,10)	0,12 (0,08)	0,04 (0,02)
Investment/Employment	10,33 (13,96)	26,38 (57,03)	16,06 (7,82)	14,81 (25,78)	15,06 (13,58)	0,25 (2,88)	13,04 (29,34)	39,95 (59,33)	26,91 (14,97)
Observations	215	55		231	34		223	16	



to pre-treatment years. Cohort 1’s phase 3 estimate is similar but statistically insignificant.

The estimates for the other outcome variables show little effect of the regulation. We can neither establish a statistically significant effect on profitability or investments in this estimation.

**Table 2:** *Results from the matched-TWFE regression.*

	Employment		Profit Margin		Investment Ratio	
Cohort 1 × Phase 1	-0.078** (0.033)	-0.077** (0.033)	0.002 (0.046)	0.001 (0.046)	-0.002 (0.052)	-0.002 (0.052)
Cohort 1 × Phase 2	-0.088* (0.050)	-0.088* (0.050)	0.026 (0.098)	0.025 (0.098)	0.009 (0.047)	0.010 (0.047)
Cohort 1 × Phase 3	-0.092 (0.074)	-0.094 (0.073)	-0.046 (0.170)	-0.047 (0.170)	0.003 (0.050)	0.004 (0.050)
Cohort 2 × Phase 2	-0.048 (0.058)	-0.047 (0.058)	-0.095 (0.104)	-0.097 (0.104)	0.067 (0.063)	0.068 (0.063)
Cohort 2 × Phase 3	0.008 (0.070)	0.005 (0.070)	-0.133 (0.152)	-0.136 (0.152)	-0.014 (0.069)	-0.013 (0.068)
Cohort 3 × Phase 3	0.101 (0.094)	0.096 (0.094)	0.232 (0.175)	0.233 (0.174)	-0.114 (0.075)	-0.113 (0.075)
Firm FEs	Yes	Yes	Yes	Yes	Yes	Yes
Phase FEs	Yes	No	Yes	No	Yes	No
Year FEs	No	Yes	No	Yes	No	Yes
Observations	6,273	6,273	6,249	6,249	6,185	6,185
Adjusted R2	0.902	0.903	0.655	0.655	0.114	0.114

The dependent variables are the log of the number of employees, the gross profit margin, and the investment to turnover ratio. The margin and the ratio are normalized on the industry level as in (4.1). Standard errors are clustered at the firm level and reported in brackets. Stars refer to \*:  $p < 0.10$ , \*\*:  $p < 0.05$ , \*\*\*:  $p < 0.01$ .

## 5.2 CS estimation

The cohort-phase estimates for the less parametric estimation from Equation 4.5 are presented in Table 3. For employment, the conclusions differ somewhat. Cohort 1 firms still reduce employment in phase 1, even though the effect is now only significant at the 10% level, but for phase 2 and 3 there are no statistically significant effects from treatment anymore. Cohort 2 phase 2’s estimate is now statistically significant, indicating a reduction in employment of about 7%. For the other cohort-phase combinations the estimates are statistically insignificant, as before.

**Table 3:** *Results from the less parametric DiD method.*

	Employment	GP margin	Investment ratio
Cohort 1 × Phase 1	-0.058* (-0.122,0.003)	0.119 (-0.262,0.453)	-0.318** (-0.508,-0.090)
Cohort 1 × Phase 2	0.034 (-0.075,0.111)	-0.187 (-0.528,0.533)	-0.466** (-0.733,-0.142)
Cohort 1 × Phase 3	0.001* (-0.106,0.074)	0.090 (-0.351,0.845)	-0.325** (-0.517,0.006)
Cohort 2 × Phase 2	-0.070 (-0.123,-0.021)	-0.005 (-0.192,0.260)	0.036 (-0.108,0.165)
Cohort 2 × Phase 3	0.002 (-0.209,0.076)	0.057 (-0.217,0.266)	-0.012 (-0.118,0.098)
Cohort 3 × Phase 3	0.069* (-0.158,0.342)	0.449*** (0.082,1.430)	-0.198* (-0.454,0.005)
Number of firms	1012	1012	1012
Regulated firms.	113	113	113
Pre-treatment	0.001 (-0.010,0.011)	-0.004 (-0.047,0.047)	0.013* (-0.003,0.049)

The dependent variables are the log of the number of employees, the gross profit margin, and the investment to turnover ratio. The margin and the ratio are normalized on the industry level as in (4.1). Bootstrapped 95% confidence intervals are in brackets. The pre-treatment estimate pools all placebo pre-treatment estimates. Stars indicate if 0 contained in confidence interval: \*: 90%, \*\*: 95%, \*\*\*: 99%

Further, cohort 3 firms increase profit margins in phase 3 with 0.45 industry-level standard deviations. The other cohorts experience no changes in their profit margins due to ETS regulation.

For investments, the effects are predominantly negative. Cohort 1 (cohort 3) firms reduce their investments-to-turnover ratio by around 0.3 (0.2) industry-level standard deviations, while cohort 2 firms do not statistically significantly respond to ETS regulation with their investment ratio. This is the sharpest deviation between the results of the TWFE and the CS estimates. Reductions in investments are inconsistent with the Porter hypothesis, but together with the negative effects on employment, could indicate a downsizing of EU activities in these firms.

### 5.3 Differences between TWFE and CS

The main differences between the results of the two estimators lie in the estimates of the effect on investments for cohort 1 and in the employment effect for cohort 1 in phases 2 and 3. Besides this, most coefficients align and mostly vary in terms of uncertainty.

The most likely reason for this deviation lies in the difference of the control groups that underlie both estimations. As outlined above, one problem with the TWFE estimator is that a treated cohort's outcomes are compared to every control firm's outcome. The matching links are lost in the regression step and thus the comparison is not between the most similar firms. The CS method compares cohorts with their own specific control firms. In the TWFE setting cohort 1-phase 1 observations could be compared to control firms that were matched to cohort 3. As cohort 1 and 3 differ, their matches likely also differ, undermining the quality of the comparison.

To make sure that the CS results are indeed based on a more specific and more comparable control group, the following section we will estimate different specifications in which we vary the underlying observations of the CS estimator.

## 6 Discussion

This section presents a discussion on the underlying assumptions and tests the robustness of our results to violations of these. By doing so, we also test how robust our results are to changes in the underlying control group. We also present results for slightly different dependent variables, and discuss our relation to the results in the literature. The presented results will in most cases be purely based on the CS estimation, as this is our preferred specification, but the TWFE results do not give additional insights above the presented and discussed ones.

## 6.1 Parallel trends

All DiD estimates rely on the parallel trends assumption to hold. Even though there is no formal test for parallel trends, one can perform some checks. We run placebo tests to see if pre-treatment periods experience differences across treatment and control groups. If so, there would be some evidence that the trends of the treatment and (matched) control would have not run in parallel in absence of the treatment.

For the TWFE method, we bring the treatment date 1, 2 and 3 years forward, drop all estimates after the actual treatment date, and see if the treatment coefficients result are statistically significant. [Figure 5](#) presents the results of these tests. For none of the date-shifts the placebo test results in statistically significant findings, implying that we cannot reject the null hypothesis of parallel trends.

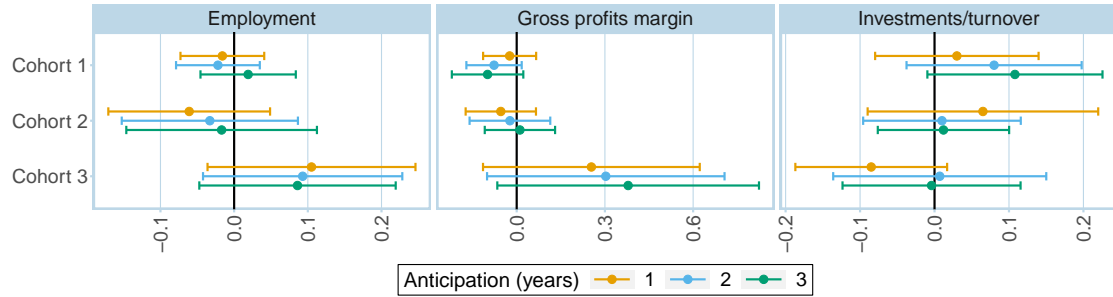
For the less parametric DiD method, we perform two similar checks. The ATTs can be estimated individually for each of the pre-treatment years, by always choosing the baseline to be the year before. We then use a Wald test to test the joint significance of these estimates; we do not reject the null for any of the three variables. We also precede in our cohort-phase setup, by aggregating all pre-treatment estimates and constructing a confidence interval in the same way as for our aggregated estimates. These results can be found in the first row of [Table 3](#). They always contain the zero in the 95% confidence interval as well.

The parallel trends tests were rejected for the analysis with the EBITDA margin. Therefore the main analysis focuses on the related gross profit margin, for which the parallel trends assumption is not rejected.

To test the robustness of the results, we adapt the matching to make the assumption more likely to hold. We have done this by using the trends, along-side the regular values, of the matching variables in the matching steps, both for the TWFE method as well as for the CS method. The resulting estimates can be found in [Figure 6](#). The results for investment completely align; for employment the coefficients are now more negative, especially for cohort 3, and similarly gross profits turn negative for cohort 3 and cohort 2-phase 2. The overall picture, however, remains the same, indicating a reduction in investments for cohort 1 and a reduction in employment for cohorts 1 and 2. The results for cohort 3 (the smallest cohort) seem to be rather volatile.

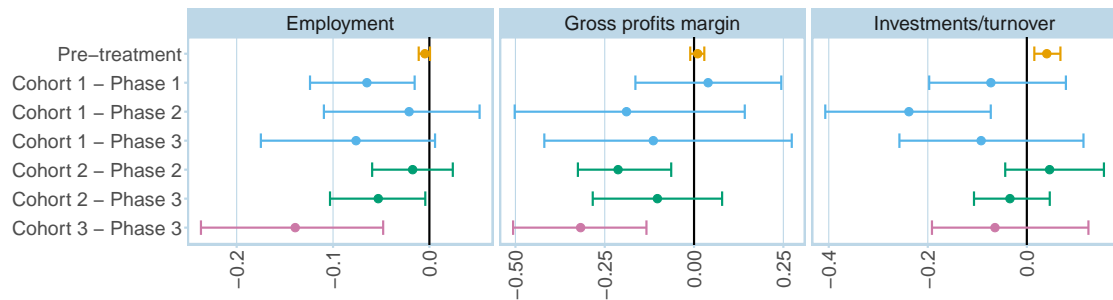
## 6.2 SUTVA

The Stable Unit Treatment Value Assumption (SUTVA) is the second necessary assumption to identification in a DiD setting. It in essence implies no spillovers between firms across treatment status. As our analysis is on the firm level, a large source of spillovers, namely that between plants in the same firm, is accounted for



**Figure 5:** *Placebo tests for parallel trends in the TWFE model.*

*Note: Placebo tests for testing the pre-treatment trend for each cohort. The colors refer to the number of years that the treatment is brought forward. E.g. 1 year of anticipation for cohort 1 tests whether there is a treatment effect in 2004. Whiskers indicate 95% confidence intervals.*



**Figure 6:** *EU ETS effects when matching on pre-trend.*

*Note: Cohort-phase coefficient estimates when the trend of the dependent variable is included in covariates to strengthen common trends. Whiskers indicate bootstrapped 95% confidence intervals. Colors refer to the different cohorts.*

already.

As regulation is on the plant level and our outcome variables are on the firm level, not all activities of the firm are regulated. This likely biases the estimates towards zero. Note that this is not a drawback per se, as this is simply how the ETS functions. The estimates are accurately representing the effect of ETS regulation on firm performance. If interested in the question what happens if all emissions were regulated, the estimates are biased towards zero.

On the other hand, there might be positive spillovers between firms through the markets that they operate in. As competition is relative, one firm's hardship can be another firm's opportunity. This can both be on the output market, as unregulated firms obtain a relative cost advantage from not being regulated, as well as on the input markets, as potential downsizing of regulated firms allows unregulated firms to snatch up employees or input supplies. This would inflate the estimates.

As both these biases to some extent relate to treatment stringency, one could compare the estimates for the different phases in the data. From [Figure 2](#) and [Figure 4](#) it becomes clear that later phases are more costly to regulated firms. [Figure 1](#) also shows that more installations of the same owner are regulated in phase 3. Estimates for the later phases should therefore suffer less from the bias towards zero, as more emissions of the firm are regulated, and more from the bias away from zero, as the relative disadvantage from regulation is exacerbated. If these biases exist, in both cases they should result in larger estimates, in absolute terms, for the later phases. [Table 3](#) does not provide evidence for either bias, as estimates for later phases within the same cohort are not further away from zero.

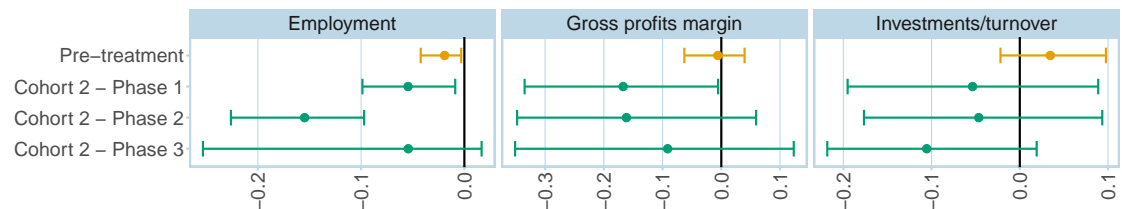
### 6.2.1 Treatment anticipation

As explained in [section 3](#), the main reason for why there are so many firms that are only regulated in the second phase is that the Dutch government excluded many firms from regulation in the first phase. As these exemptions are public information, it seems likely that these firms expected to be regulated in phase 2. If so, the firms in cohort 2 would have already anticipated treatment in 2003, which would violate our assumptions on the anticipation, and would also make these firms an improper control for cohort 1 in phase 1.

On the other hand, this might enable us to roughly disentangle an anticipation from an actual treatment effect, by treating cohort 2 as already being regulated in phase 1. The estimate from such an experiment also provides cohort 2's anticipation effect of being regulated in phase 1, and provides an adjusted estimate of the effect in phases 2 and 3, with an adjusted control group and base year (2003).

The results for this estimation can be found in [Figure 7](#). The results somewhat align with the main results for phases 2 and 3, although the statistical significance

differs. Interestingly the negative employment effect in phase 2 was preceded by a small negative effect in phase 1 already, which could imply that firms did anticipate the regulation and started reducing their employment in anticipation. The same can be concluded for the gross profit margin.



**Figure 7:** *EU ETS effects when treating cohort 2 as of Phase 1.*

*Note:* Cohort-phase coefficient estimates for cohort 1 when allowing it to show a regulation response as of Phase 1. Whiskers indicate bootstrapped 95% confidence intervals. Colors refer to the different cohorts.

## 6.2.2 Compare to never-treated firms

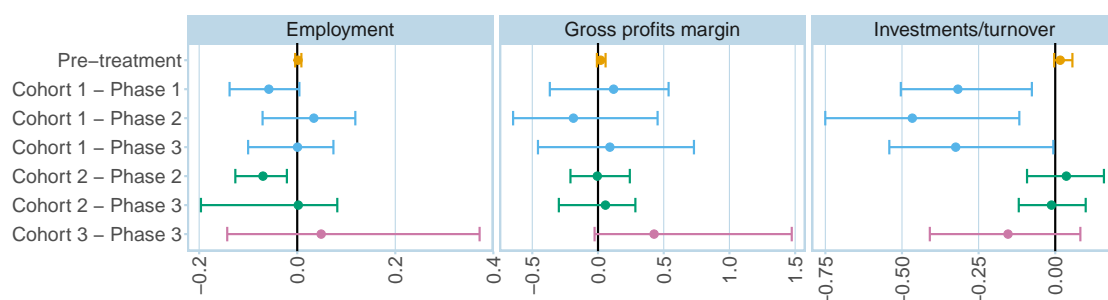
To further alleviate concerns about potential anticipation effects in the control group, we can rid the control group of firms that in later phases become regulated. The control group then exclusively consists of firms that are never regulated under the EU ETS. The advantage is that the control group is then less likely to contain anticipation effects to upcoming regulation, as was indicated in the previous subsection. The disadvantage is that restricting the set of available controls makes the matches poorer.

Figure 8 shows the results of this exercise. Most notable is the close similarity of the results to our main analysis, making us further confident in our estimates.

## 6.3 Alternative measures as dependent variables

To test if our results rely on our choice of the dependent variable, we test if (1) the scaling of investment and (2) the choice of the profit margin is important for our results. We thus first scale investments (still into all fixed assets) by the number of employees in full-time equivalents. The results are presented in Figure 9. The results are in line with the main specification, in which we scaled investments by turnover.

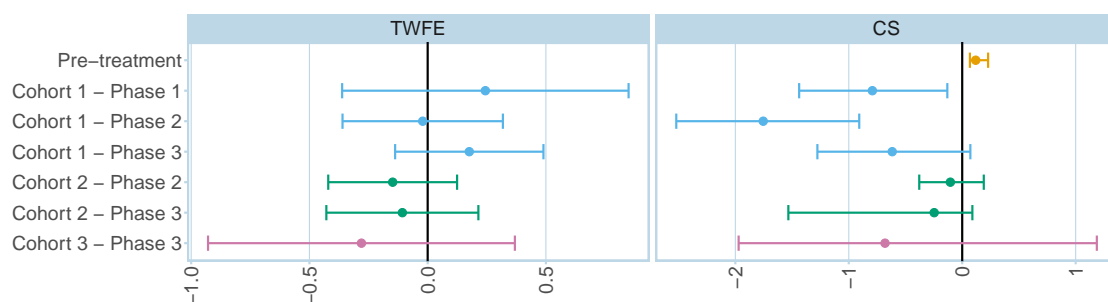
Second, we look at the EBITDA margin instead of the gross profit margin. EBITDA is often used as a more harmonized measure of earnings, as it takes the gross profits before the financial result. As mentioned earlier the EBITDA margin does not pass the placebo tests and hence we have rely on the gross profit margin



**Figure 8:** *EU ETS effects when strictly comparing to never treated control firms.*

*Note: Cohort-phase coefficient estimates for cohorts 1-3 when strictly allowing for comparisons with never-treated control firms. Whiskers indicate bootstrapped 95% confidence intervals. Colors refer to the different cohorts.*

instead. For completeness the results for the EBITDA margin can be found in [Appendix D, Figure D.4](#), and have similar signs as that of the gross profit margin, with, however, smaller confidence intervals.



**Figure 9:** *EU ETS effects on investments over employment.*

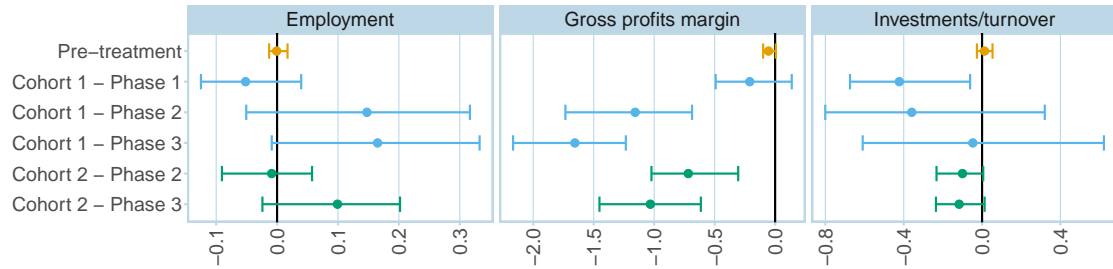
*Note: Cohort-phase coefficient estimates for cohorts 1-3 for investments scaled by the number of employees (FTE). Both the TWFE and CS results are provided. Whiskers indicate 95% confidence intervals, which for CS are bootstrapped. Colors refer to the different cohorts.*

## 6.4 Fully balanced panel

To see in how far our results are sensitive to firms exiting our sample at a later stage, we redo the estimation on a fully balanced panel. This greatly reduces the number of observations and firms, which also prevents us from reporting cohort 3 findings due to privacy requirements from CBS.



The results are presented in Figure 10 and show that while the results for investments are still in line with the main results, those for the profit margin are quite different. The results indicate that on this smaller panel, the regulated firms experience large losses in profitability. As this is based on few observations, we assume that these results are based on few outliers, but we take this as further evidence that the profitability results are rather volatile and inconclusive. For employment our results confirm the early reduction in employment in phase 1, but also show that this effect was likely temporary.



**Figure 10:** *EU ETS effects when enforcing a balanced panel.*

*Note: Cohort-phase coefficient estimates for cohorts 1-3 when enforcing a balanced panel, with thus only 135 underlying firms (52 treated). Each firm in the sample is observed for all the years in the sample. Whiskers indicate bootstrapped 95% confidence intervals. Colors refer to the different cohorts.*

## 6.5 Comparison to literature

Most articles that have studied the ETS's effect have found little side effects from the regulation. In a sense our study's overall findings are not too different from the literature's findings. We do not find conclusive indications of a reduction in profit margins and the findings of negative employment effects seem temporary. However, we do find that cohort 1 reduces investments in all phases and employment in phase 1. We have shown that cohort 1 mostly consists of the most energy-intensive firms, which is likely linked to this result. Other studies do not reach these conclusions.

The reason for the negative effects could be based in the high export-orientation of the Dutch economy. The exposure to trade might have incentivized regulated firms to downscale their operations within the Netherlands as a response to the regulation, even before the regulation became more stringent. Such anticipation of future strictness of the regulation would also be in line with the anticipation effect that we find for cohort 2 firms. By disentangling the effects for the very energy-intensive firms in cohort 1 from the later, less energy-intensive cohorts, we show that heterogeneity between the cohorts plays an important role. Not all firms will

respond the same to the EU ETS. The fact that the negative employment effect is not maintained in the last phases, however, suggests that some effects might have been temporary.

Like Löschel et al. (2019) we also find statistically significant effects rather in the less-parametric DiD setting than in the TWFE regression, even though our results differ. This highlights the importance of choosing the right estimation methods when studying the ETS'effects.

## 7 Conclusion

This paper studies the effects of the EU ETS on the competitiveness and the investment behavior of regulated manufacturing firms in the Netherlands. Motivated by large differences in energy intensity between firms and by an increase in regulatory stringency over time, we pay special attention to treatment effect heterogeneity. We employ two difference-in-differences designs to estimate the average treatment effect on the treated. As a benchmark, we employ a classical matched two-way fixed effects regression. To better allow for heterogeneity, we also employ a recent, more flexible DiD method introduced by Callaway and Sant'Anna (2021).

We estimate the effects of EU ETS regulation on employment, the gross profit margin and investments over turnover. To capture the heterogeneities we estimate the effect of phases 1-3 in the EU ETS for each cohort, whereby a firm belongs to one of the three cohorts if it was first regulated in that respective phase. We make use of public data from the European Union Transaction Log and restricted-access microdata from Statistics Netherlands.

Our matched TWFE approach results in some evidence of employment loss in cohort 1 firms. For phase 1 and 2 these firms lose around 8% of employment compared to the non-regulated baseline. For profits and investments the TWFE method does not yield statistically significant findings.

In the preferred and more flexible DiD design, we find that cohort 1 firms reduce employment by 6%, but this finding is only significant at the 10% level and does not persist in later phases. Cohort 2 firms see a 7% decrease in employment in phase 2. For the other cohort-phase combinations the results are not statistically significant. The results with the profit margin as dependent variable are inconclusive. Cohort 1 firms reduce investments across all phases. The other cohorts do not seem to respond to regulation with changes in investments.

Taken together, our findings point to some employment loss in the Dutch manufacturing industry due to the EU ETS regulation. This is most notable in the first regulation phases for cohort 1 and 2. In phase 3 employment is restored to baseline levels for both cohorts. The temporary employment loss is between 7-9% in our main TWFE and flexible DiD estimations. A priori we expected such an

effect to be higher in later phases, when treatment stringency increased, but the result might indicate that firms reacted early, which is also consistent with some evidence for anticipation effects that we find by studying later treated firms.

We further find a negative effect on investments for cohort 1 firms, especially in phase 1 and 2, but only when applying the more flexible DiD method. These firms lose between 0.32 and 0.47 industry-level standard deviations in their investments to turnover ratio. We can conclude that a Porter Hypothesis style boost to investments in response to environmental regulation seems unlikely. For profitability the results are mixed and often statistically insignificant.

We thoroughly test the two main underlying assumptions of the DiD methodology, namely that of common trends and SUTVA. Common trends are tested for with pre-treatment placebo tests. We have only interpreted the findings of estimations for which the common trend hypothesis is not rejected. Matching on trends for the matching variables also does not change the main conclusions much, except that the later cohorts reduce employment and profits more. We have also extensively discussed SUTVA and can conclude that spillovers play a small role, but that anticipation to treatment might exist.

Our results fit into the literature in two ways. First, the different findings between our matched TWFE method and the more flexible DiD method highlight the importance of the right DiD design and estimator, as heavily discussed in recent literature. The most important difference here is how the counterfactual is composed and constructed. Second, our findings add to the debate on negative and positive side-effects of environmental regulation. Using data up to the end of phase 3 (2020) allows studying heterogeneity over time. We conclude that some worries over employment loss might be warranted by our findings, but they are of temporary nature. Profits seem mostly unaffected, but investments seem to decrease for the most energy-intense firms when regulated.

Research of the EU ETS's effects remains of interest, as longer time series allow for the evaluation of medium and long-term effects. This can be informative to policy makers that consider the implementation or strengthening of environmental policy. Future research will also allow for the analysis of changes in regulatory stringency, which we here already exploited to some extent when discussing SUTVA. Analysis of phase 4 reforms and the high EUA prices as of 2021 might provide new insights.

## References

- Abadie, A. (2005). Semiparametric Difference-in-Differences Estimators. *The Review of Economic Studies*, 72(1), 1–19.
- Abrell, J. (2021). Database for the European Union Transaction Log [Database, available at <https://euets.info>].
- Borghesi, S., Cainelli, G., & Mazzanti, M. (2015). Linking emission trading to environmental innovation: Evidence from the Italian manufacturing industry. *Research Policy*, 44(3), 669–683.
- Calel, R., & Dechezleprêtre, A. (2016). Environmental policy and directed technological change: Evidence from the European carbon market. *Review of Economics and Statistics*, 98(1), 173–191.
- Callaway, B., & Sant’Anna, P. H. (2021). Difference-in-differences with multiple time periods [Themed Issue: Treatment Effect 1]. *Journal of Econometrics*, 225(2), 200–230.
- Colmer, J., Martin, R., Muûls, M., & Wagner, U. J. (2022). *Does pricing carbon mitigate climate change? Firm-level evidence from the European Union Emissions Trading Scheme* (Discussion Paper No. DP16982). CEPR.
- Daw, J. R., & Hatfield, L. A. (2018). Matching and regression to the mean in difference-in-differences analysis. *Health services research*, 53(6), 4138–4156.
- de Chaisemartin, C., & D’Haultfoeuille, X. (2022). *Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey* (Working Paper No. 29691). National Bureau of Economic Research.
- Dechezleprêtre, A., Gennaioli, C., Martin, R., Muûls, M., & Stoerk, T. (2019). *Searching for carbon leaks in multinational companies* (Discussion Paper No. 1601). Centre for Economic Performance, LSE.
- Ellerman, A. D., & Buchner, B. K. (2008). Over-allocation or abatement? A preliminary analysis of the EU ETS based on the 2005–06 emissions data. *Environmental and Resource Economics*, 41(2), 267–287.
- European Commission. (2004). Commission decision of 29 October 2004 concerning the temporary exclusion of certain installations by the Netherlands from the Community emissions trading scheme pursuant to Article 27 of Directive 2003/87/EC of the European Parliament and of the Council [Decision C(2004)4240-3].
- European Commission. (2005). Commission decision of 22 March 2005 concerning the temporary exclusion of certain installations by the Netherlands from the Community emissions trading scheme pursuant to article 27 of directive 2003/87/ec of the European Parliament and of the Council [Decision C(2005) 866 final].
- European Parliament, Council of the European Union. (2003). Establishing a scheme for greenhouse gas emission allowance trading within the community [Directive 2003/87/EC].
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing [Themed Issue: Treatment Effect 1]. *Journal of Econometrics*, 225(2), 254–277.

- Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies*, 64(4), 605–654.
- Hintermann, B. (2010). Allowance price drivers in the first phase of the EU ETS. *Journal of Environmental Economics and Management*, 59(1), 43–56.
- Jaraite-Kažukauske, J., & Di Maria, C. (2016). Did the EU ETS make a difference? an empirical assessment using Lithuanian firm-level data. *The Energy Journal*, 37(1).
- Klemetsen, M., Rosendahl, K. E., & Jakobsen, A. L. (2020). The impacts of the EU ETS on Norwegian plants' environmental and economic performance. *Climate Change Economics*, 11(01).
- Löschel, A., Lutz, B. J., & Managi, S. (2019). The impacts of the EU ETS on efficiency and economic performance – An empirical analyses for German manufacturing firms. *Resource and Energy Economics*, 56, 71–95.
- Marin, G., Marino, M., & Pellegrin, C. (2018). The impact of the European Emission Trading Scheme on multiple measures of economic performance. *Environmental and Resource Economics*, 71(2), 551–582.
- Martin, R., Muûls, M., De Preux, L. B., & Wagner, U. J. (2014). Industry compensation under relocation risk: A firm-level analysis of the EU emissions trading scheme. *American Economic Review*, 104(8), 2482–2508.
- Petrick, S., & Wagner, U. J. (2014). *The impact of carbon trading on industry: Evidence from German manufacturing firms* (Working paper No. 2389800). Available at SSRN.
- Porter, M. E., & Van der Linde, C. (1995). Toward a new conception of the environment-competitiveness relationship. *Journal of Economic Perspectives*, 9(4), 97–118.
- Sant'Anna, P. H., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1), 101–122.
- Teixidó, J., Verde, S. F., & Nicolli, F. (2019). The impact of the EU Emissions Trading System on low-carbon technological change: The empirical evidence. *Ecological Economics*, 164 (Article Number: 106347).
- Verde, S. F. (2020). The impact of the EU Emissions Trading System on competitiveness and carbon leakage: The econometric evidence. *Journal of Economic Surveys*, 34(2), 320–343.

## A Data details

### A.1 Statistics Netherlands (CBS)

The units in the CBS data are partially constructed by CBS itself. Especially the Business Unit (BE) is a construct that is generated by CBS. Here we will discuss how these units are constructed.

#### A.1.1 Business Unit (BE)

The business unit (BE) captures outward-facing (i.e. non-internal) Dutch production or service-provision that can be seen as one unit. This means that legal firm structures are grouped by purpose into BEs, e.g. a unit producing wooden furniture. This provides several advantages and disadvantages. The main advantage is that the BE is a unit structure that captures economic activity well. Legal firm structures often only exist for fiscal reasons and do not represent economic activity or choices well. The disadvantage is that BEs are constructed and that their composition can change over time, even though these changes might be representative of economic activity within the BE.

### A.2 EU ETS

For the data on the EU ETS, coming from [EUETS.INFO](http://EUETS.INFO), a few transformations are needed.

The main problem occurs when installations change owner. This event is poorly captured by the data and therefore requires manual corrections. The corrections of ownership change were done in the following steps.

1. From the European Commission's Union Registry the lists of (stationary) installations for each phase are downloaded.<sup>10</sup>
2. The owners of each installation are compared across phases. If the owners are unchanged between phases, they are assumed to have been the same within that phase.
3. For the installations of which owners have changed between phases, we search the internet for further information to determine whether there was a transfer of ownership and between whom. From sources like news articles or websites that provide ownership data, we deduce when ownership has changed and to who. Two common situations occur, namely (1) ownership of installations is transferred within a firm group, which effectively means the installation

---

<sup>10</sup>These lists can be found for Phase 1, 2 and 3 on the EC's [website](http://EUETS.INFO).

has the same ultimate owner and (2) another firm purchases the installation, sometimes because the previous owner went bankrupt.

4. For installations that saw their owner change but for which we find no information when this took place, we assumed the change to take place on the day the new phase started.

The dates of ownership change then have to be reconciled with the annual data. For this, the year was chosen in which the ownership change has taken place and this year is considered to be the year in which the new owner takes economic responsibility of the installation.

### **A.3 Details on merging the EUTL with CBS data**

Data that is imported into the CBS environment and that is identified on the chamber of commerce (in Dutch: KvK) number, like the ETS data, is encrypted on the same level. So installations under the EU ETS are imported into the CBS environment and encrypted. Encrypted chamber of commerce numbers can then be used to link EU ETS regulation to the business units.

Based on this encryption, one can find the corresponding CBS person (Dutch: persoon) in each year. This CBS person presents a layer in between the detailed KvK number and the final identifier level, business units (BEs). The CBS persoon itself is just a one to one linking from the KvK number to a CBS internal identifier. In some rare years a KvK number is assigned to two CBS persons within a year. This is because CBS draws from multiple sources which can cause duplicate links. In these cases, we have decided to assign the KvK number to the later created CBS person within that year.

The original ETS plant is thus assigned to a BEID in each year, ownership changes between years are thus uncritically represented here. However, in some years a CBS person is assigned to two BEIDs, which can happen if ownership changes within a year. In these cases, we assign the later BEID to the plant.

The CBS data sets are all identified on the BEID level and so we can in the next step merge the ETS plants to the CBS data sets. In each of these steps some of the companies cannot be assigned to another identifier or data set, such that in the end not all ETS firms can be merged. There is, however, no systemic bias in this. After consultation with CBS, the majority of the firms that we were not able to link stem from site that has merged several ETS installations under one account holder, which are then impossible to link to the BEID in our data.

## B Technicalities of estimation strategy

### B.1 Further explanation and definitions of the group-year specific ATT

We here give the definitions of the inverse probability and outcome regression adjustments as well as their underlying interpretation.

$$\hat{w}_{jc}^{treated} = \frac{G_{jc}}{\frac{1}{N} \sum_i G_{jc}} \quad (\text{B.1})$$

$$\hat{w}_{jc}^{control} = C_{jc} \frac{\frac{p_{jc}(X_j, \hat{\pi}_c)}{1-p_{jc}(X_j, \hat{\pi}_c)}}{\frac{1}{N} \sum_i \frac{p_{jc}(X_j, \hat{\pi}_c)}{1-p_{jc}(X_j, \hat{\pi}_c)}} \quad (\text{B.2})$$

with  $G_{jc}$  being a dummy for if a firm is in the respective treatment group or not,  $C_{jc}$ , a dummy that is one if the firm can serve as a control for that treatment cohort, thus incorporating never treated as well as not yet treated firms, and  $p_{jc}$  as the estimated propensity score for each firm (giving the probability of being in that treatment cohort), based on the controls and the estimated coefficients  $\hat{\pi}_c$  from a logistic regression model. This procedure thus weights controls that are more likely to be treated higher than firms that are unlikely to be treated.

$\hat{m}_{jc}(X_j, \hat{\lambda}_{ct})$  is the estimator of  $\mathbb{E}[Y_t - Y_{base} | X, C = 1]$ . It is thus the difference in predicted values between year  $t$  and the base year for the treated firms, if they were untreated. One thus runs  $y_{jt} - y_{jb} = \lambda X_j + \varepsilon_j$  only on the sample of the untreated units, to estimate the change in outcomes that can be predicted by the covariates and then uses this  $\hat{\lambda}$  to predict  $\hat{m}_{jt}(X_j, \hat{\lambda}_{tc}) = \widehat{y_{jt} - y_{jb}}$ , in this case both for the treated and untreated units.

## C Matching

Our matching algorithm for the TWFE estimation is presented in [Algorithm 1](#). The algorithm is designed to match treated firms to similar enough control firms in order to make a sensible comparison between their economic outcomes. It also attempts to filter for good data quality, e.g. by only considering firms that are observed for several consecutive years around treatment.

### Algorithm 1: Matching

1. Enforce common support between treated and control units



- (a) For each baseline year, we drop all observations that are outside the common support of the treated and control group.
2. Select treatment period
    - (a) Take treatment period  $T \in T^p$ , where  $T^p$  is the set of treatment periods, i.e. the years 2005, 2008 and 2013 for phase 1, phase 2 and phase 3 ( $p$ ) in the EU ETS respectively.
  3. Select observations to be potentially matched
    - (a) From the ever-treated EU ETS firms, select only those observations that are first regulated in phase  $p$ . Keep all observations from the never-treated group.
    - (b) Only keep units that are observed for all of the years in  $(T - pre, T + post)$ , where we set  $pre = 2$  and  $post = 3$ . This guarantees that resulting matches can be observed around the treatment period.
    - (c) Select only the observations at  $T - pre$ , dropping the panel structure. This year will be the pre-treatment matching period.
  4. Similarity scoring and match decision
    - (a) Measure the Mahalanobis distance between all observations in the selected sample across treatment status for the variables  $X^m$ .<sup>11</sup>  $X^m$  are the matching variables for which we take the number of employees, turnover, wage expenses, energy expenses, and value added and their squared values. We also restrict matches to be only within a 2-digit sector code. Matches across sectors are not allowed.
    - (b) For each treated unit collect the  $H$  closest neighbors based on the Mahalanobis distance. We opt for  $H = 5$  and we do allow for replacement. We also allow for ties, meaning ties are not randomly broken but rather all are included in the result. For the implementation of this step and the previous step we leverage on the `Matching` package's `Match` function in R.
  5. Store matching outcome
    - (a) Remaining matches are stored under matching year  $T - pre$ .

---

<sup>11</sup>The Mahalanobis distance between treated ( $T$ ) unit A's covariate vector  $x_A$  and control ( $C$ ) unit B's covariate vector  $x_B$  is given by  $d(A, B) = \sqrt{(x_A^T - \mu^T)S^{-1}(x_B^C - \mu^C)}$ , where  $S$  is the variance-covariance matrix between  $x^T$  and  $x^C$  and where the  $\mu$ s are the means of their respective series. Note that this distance measure is like a variance-corrected normalized Euclidean distance.

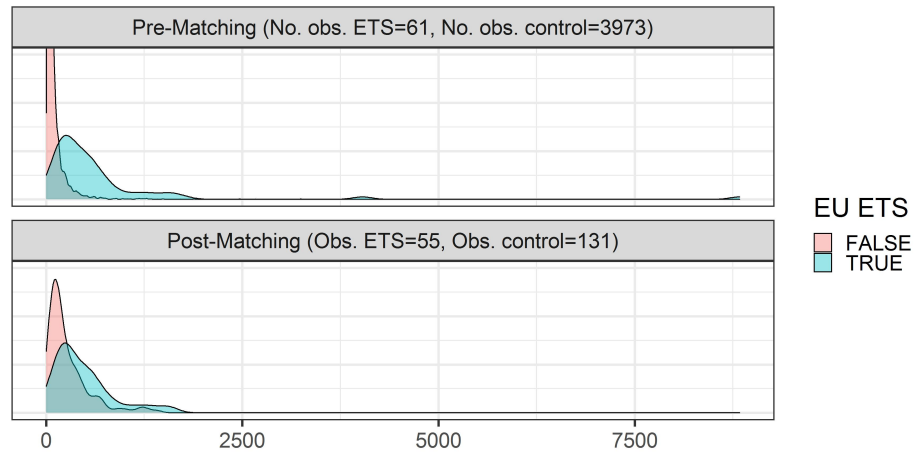
6. Next treatment period

- (a) If not all treatment periods in  $T^p$  are covered yet, select the next value in  $T^p$  and repeat the algorithm from step 2.

Table 1 provides the balancing table after matching. Figure C.1, Figure C.2 and Figure C.3 show the distributions of selected variable for regulated versus non-regulated firms before and after the matching procedure for the pre-phase 1 year 2003, pre-phase 2 year 2006 and the pre-phase 3 year 2011 respectively.

### Distributions for No. of employees

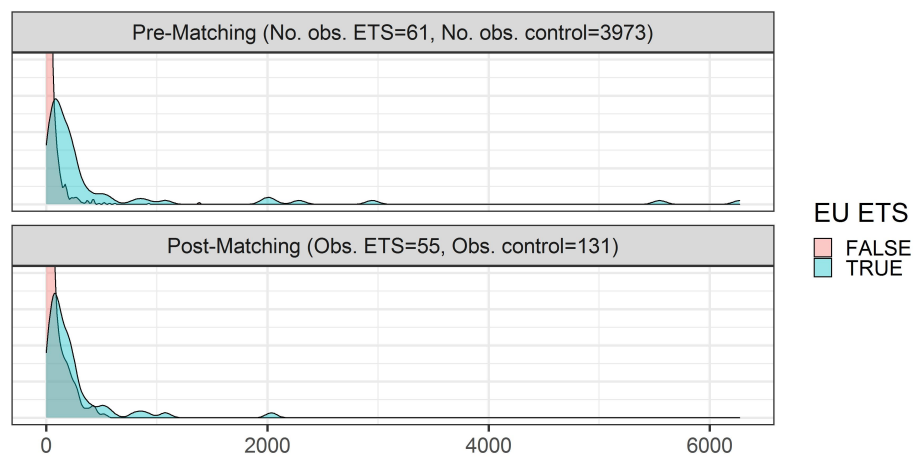
Year = 2003



(a) *Number of employees*

### Distributions for Turnover (Millions EUR)

Year = 2003

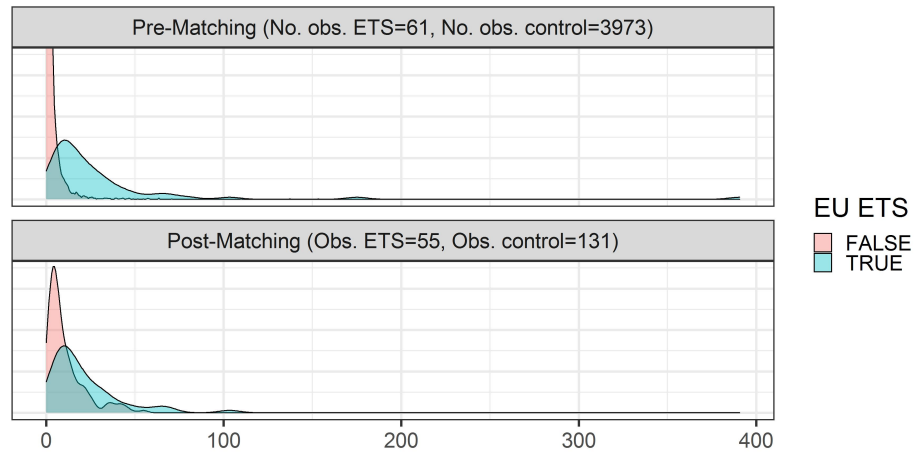


(b) *Turnover*

**Figure C.1:** *Distributions of variables before and after matching for treated and control firms in 2003.*

### Distributions for Wages (Millions EUR)

Year = 2003



(c) *Total wages*

### Distributions for Energy expenses (Millions EUR)

Year = 2003

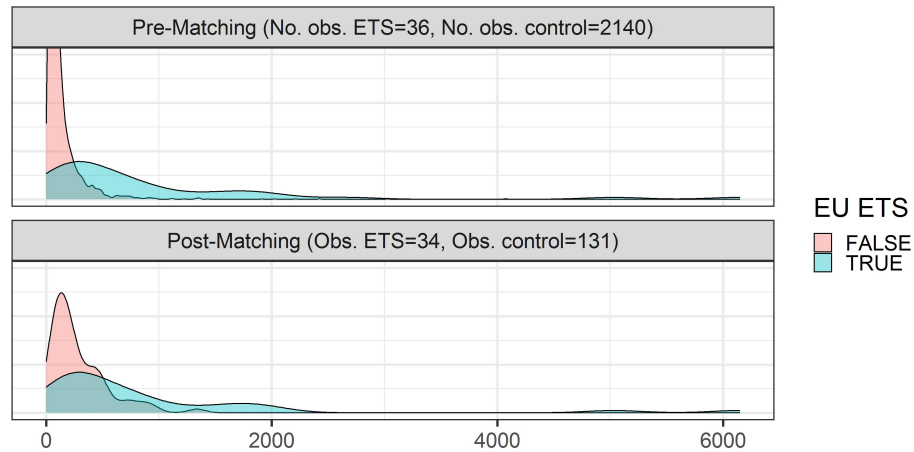


(d) *Energy expenditures*

**Figure C.1:** *Distributions of variables before and after matching for treated and control firms in 2003. (Cont'd.)*

### Distributions for No. of employees

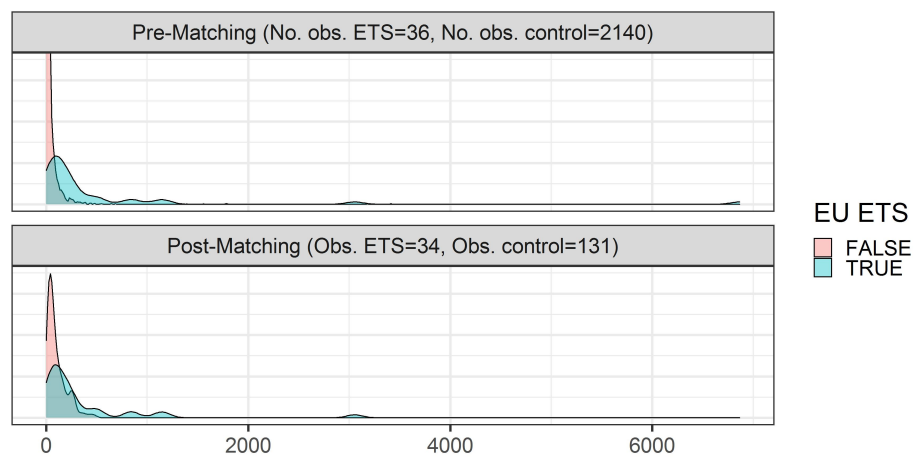
Year = 2006



(a) *Number of employees*

### Distributions for Turnover (Millions EUR)

Year = 2006

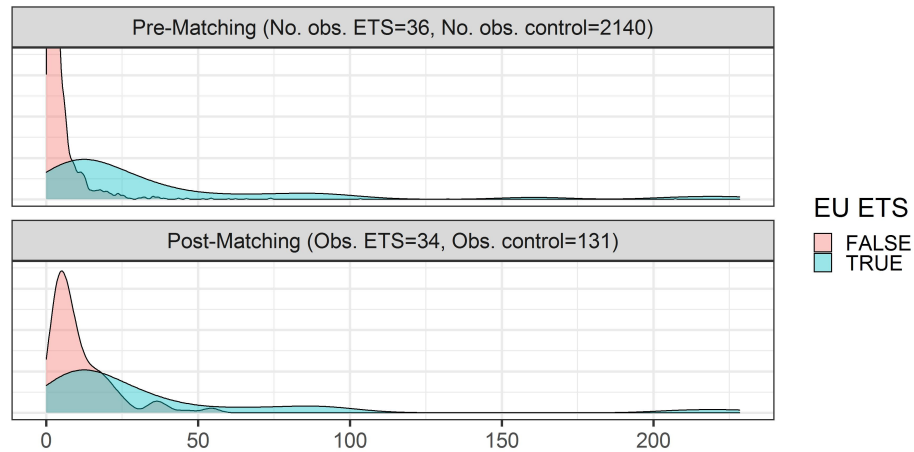


(b) *Turnover*

**Figure C.2:** *Distributions of variables before and after matching for treated and control firms in 2006.*

### Distributions for Wages (Millions EUR)

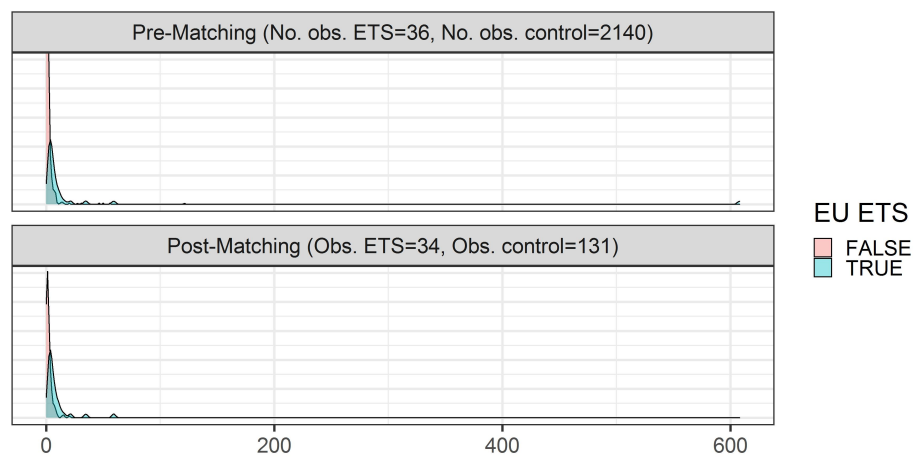
Year = 2006



(c) *Total wages*

### Distributions for Energy expenses (Millions EUR)

Year = 2006

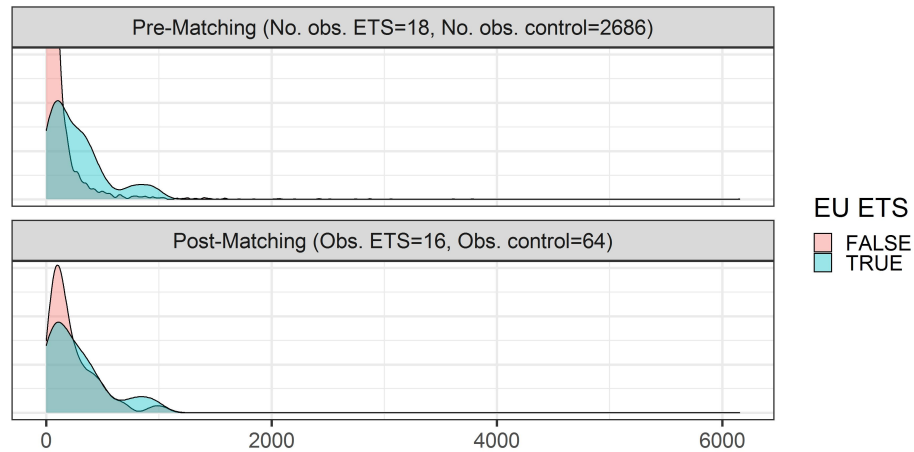


(d) *Energy expenditures*

**Figure C.2:** *Distributions of variables before and after matching for treated and control firms in 2006. (Cont'd.)*

### Distributions for No. of employees

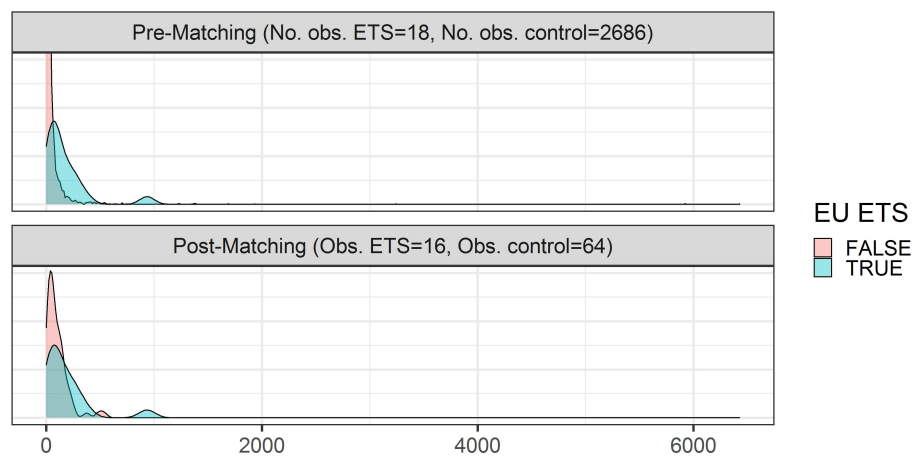
Year = 2011



(a) *Number of employees*

### Distributions for Turnover (Millions EUR)

Year = 2011

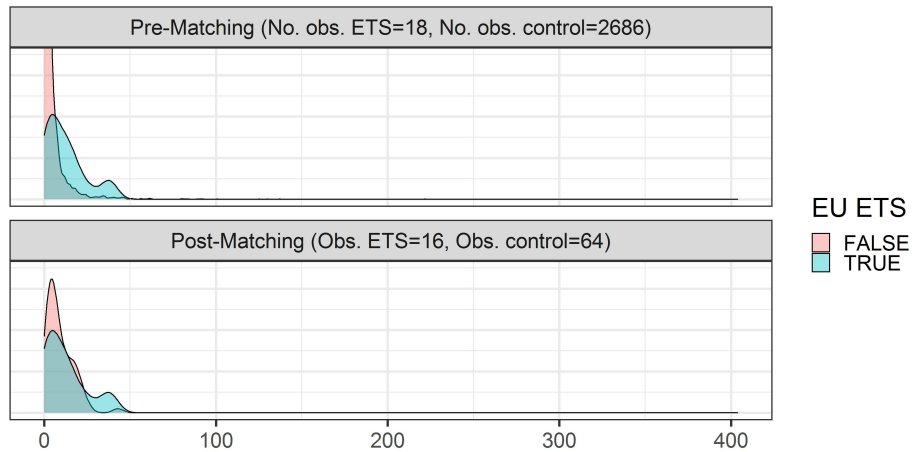


(b) *Turnover*

**Figure C.3:** *Distributions of variables before and after matching for treated and control firms in 2011.*

### Distributions for Wages (Millions EUR)

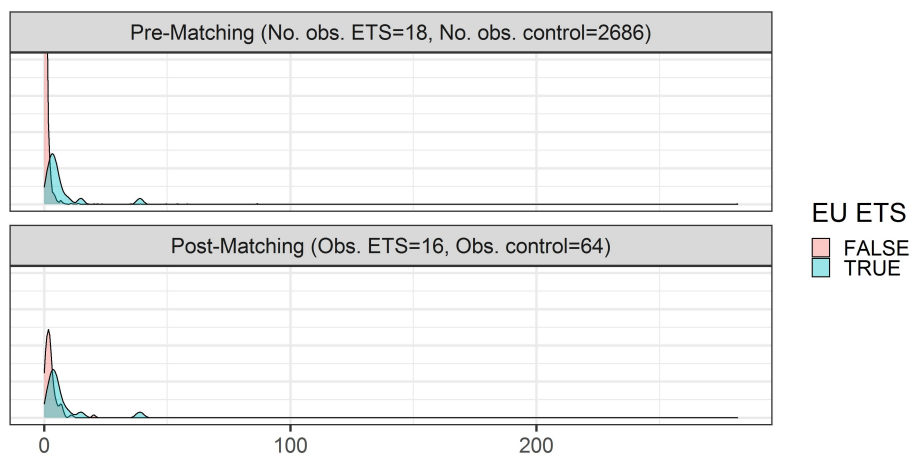
Year = 2011



(c) *Total wages*

### Distributions for Energy expenses (Millions EUR)

Year = 2011

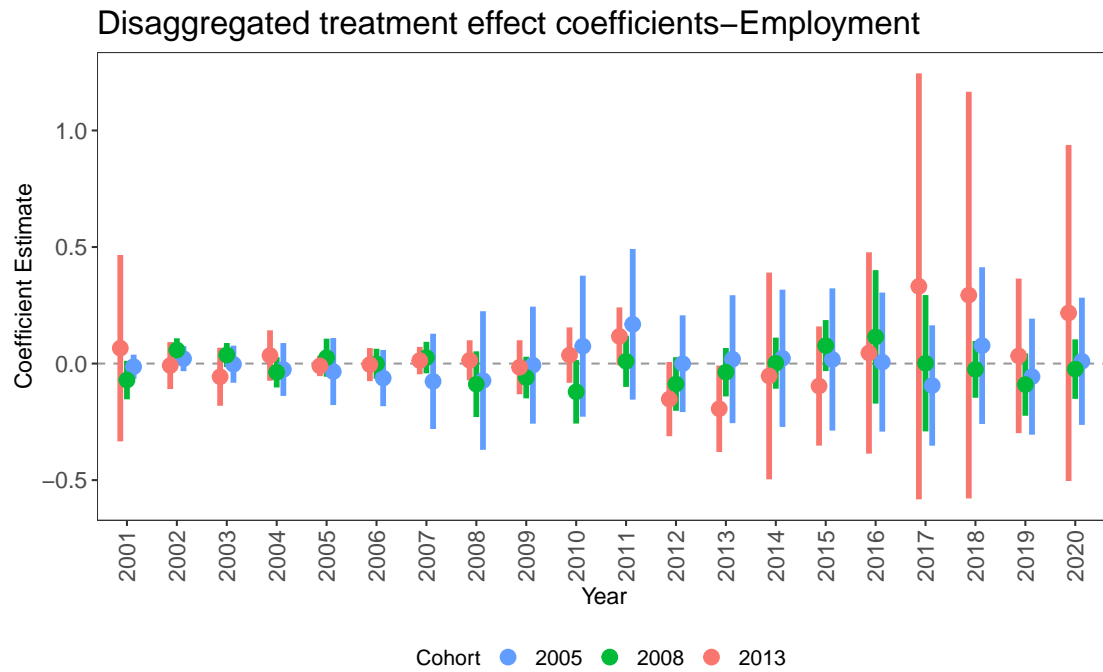


(d) *Energy expenditures*

**Figure C.3:** *Distributions of variables before and after matching for treated and control firms in 2011. (Cont'd.)*

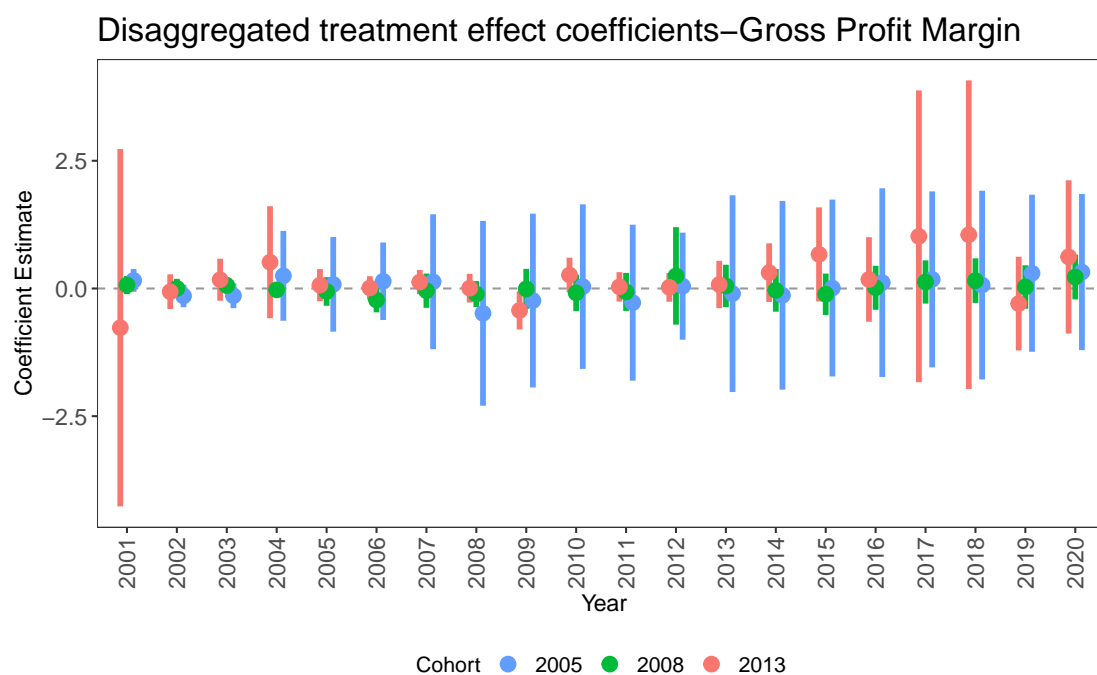


## D Additional tables and figures



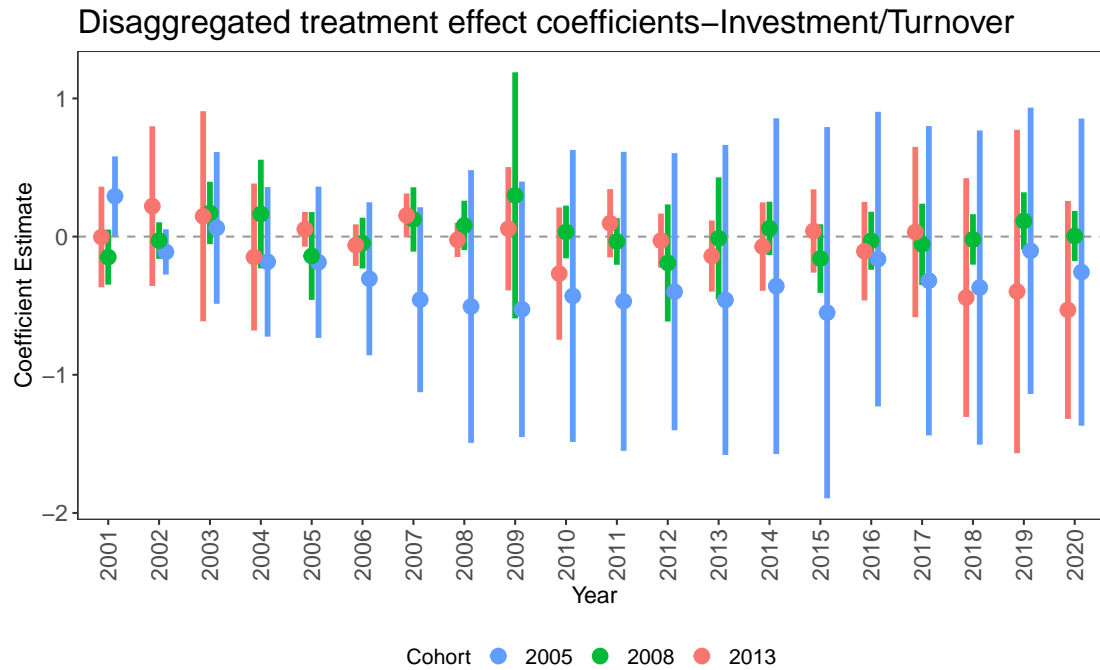
**Figure D.1:** *Treatment effect estimates for employment.*

*Note:* Non-aggregated coefficient estimates for employment from analysis of [Equation 4.3](#). Bars present 95% confidence regions. Pre-treatment estimates are placebo estimates.



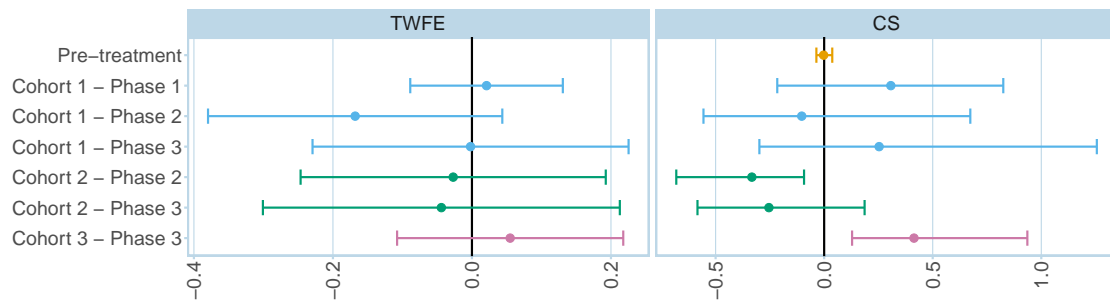
**Figure D.2:** *Treatment effect estimates for employment.*

*Note:* Non-aggregated coefficient estimates for gross profit margin from analysis of [Equation 4.3](#). Bars present 95% confidence regions. Pre-treatment estimates are placebo estimates.



**Figure D.3:** *Treatment effect estimates for employment.*

*Note: Non-aggregated coefficient estimates for turnover from analysis of Equation 4.3. Bars present 95% confidence regions. Pre-treatment estimates are placebo estimates.*



**Figure D.4:** *EU ETS effects on EBITDA margin.*

*Note: Cohort-phase coefficient estimates for cohorts 1-3 for EBITDA margin. Both the TWFE and CS results are provided. Whiskers indicate 95% confidence intervals, which for CS are bootstrapped. Colors refer to the different cohorts.*