Analysing the Effectiveness of International Environmental Policies: The Case of the Kyoto Protocol*

Christian Almer† and Ralph Winkler‡

June 29, 2015

Abstract

We study the effectiveness of emission targets under the Kyoto Protocol with respect to reducing CO_2 emissions. Using country-level and US state-level panel data and employing the synthetic control method, we do not find a significant effect for any of the major emitters among the Annex B countries with binding emission targets. We also show that – in general – evaluating the effectiveness of international environmental policies at the country level comes with a number of empirical challenges that may invalidate findings based on more traditional panel data approaches.

Keywords: Climate Policy, International Environmental Agreements, Kyoto Protocol, Synthetic Control Method

JEL-Classification: K33, Q54

*We would like to thank Alberto Abadie, Philip Cooper, Timo Goeschl, Michael Greenstone, Steve Stillman, Ulrich Wagner and seminar/conference participants at Bath, Chicago, and Lucerne for invaluable comments on an earlier draft of the paper.

†Department of Economics and Institute for Policy Research, University of Bath, 3 East, Bath BA2 7AY, UK; c.almer@bath.ac.uk

‡Department of Economics and Oeschger Centre for Climate Change Research, University of Bern, Schanzeneckstrasse 1, CH-3012 Bern, Switzerland; mail@ralph-winkler.de

1 Introduction

The Kyoto Protocol (KP) is – to date – the only international climate policy with binding greenhouse gas (GHG) emission targets for at least some countries. However, the initial commitment period ranging from 2008–2012 has already expired and no agreement on a follow-up treaty has been reached so far.[1] In fact, for almost a decade we have been witnessing the struggle of the international community to agree on a successor of the KP. This tenacious process is no surprise given the rich theoretical literature on climate treaties and its disillusioning findings regarding the likelihood of a global agreement (see, for example, Finus, 2008 for a review of this literature). In addition, the KP has been heavily criticized since its emergence in 1997.[2] Yet, the international community persists on the idea of a global follow-up agreement with emission targets for a broader set of countries that is hoped to be agreed on during the COP 21 conference in Paris in December 2015.

In contrast to the theoretical economic literature and the widespread criticism mentioned above, this endeavor is supported by a recent and increasing body of empirical literature on the effectiveness of the KP. Aichele and Felbermayr (2012), Aichele and Felbermayr (2013) and Grunewald and Martinez-Zarzoso (2015) consistently find that countries with binding emission targets under the KP have lower CO_2 emissions than they would have had in the absence of these targets. In fact, they estimate a statistically significant average CO_2 reduction effect of 7 to 10%.

In this paper, we argue that the statistically significant treatment effect is spurious and likely to be caused by a misspecification of the empirical model. We test for the existence of observable emission reductions in 15 major Western Annex B countries with binding GHG emission targets under the KP. We show that opposing trends in CO_2 emissions between countries with (Annex B countries) and without (non-Annex B countries) binding emission targets under the KP lead to a violation of the common trend assumption which is critical in this particular setting (difference-in-differences or extensions of it). Not addressing the

[1] A second commitment period of the KP was proposed in 2012, known as the Doha Amendment. However, several countries that participated in the first commitment period have withdrawn their support.

[2] In December 1997 The Economist (1997) already prognosticated that the USA will never be able to ratify the KP, as it would never be approved by the U.S. Senate. Prins and Rayner (2007) criticize its inflexible top-down architecture, which had been borrowed from past international treaties regulating chlorofluorocarbons, sulphur emissions and nuclear weapons, and "was always the wrong tool for the nature of the job." Also the economics profession found little praise for the KP. While Barrett (1998) argued from a political economy point of view that the KP hardly deters non-participation and non-compliance, Copeland and Taylor (2005) criticize that its design neglects important lessons from trade theory. Other authors animadvert the level of the emission targets (e.g., Tol, 2000) or discuss the challenges of the flexibility mechanisms (Zhang and Wang, 2011).

opposing trends results in highly significant treatment effects which vanish entirely once the pre-treatment trends have been balanced. We explicitly address the common trend assumption (and selection into treatment) by employing the synthetic control method (Abadie and Gardeazabal, 2003; Abadie et al., 2010, 2015).

The main obstacle to analyzing the effect of the KP on the emissions of Annex B countries is the identification of the correct counterfactual, i.e. business-as-usual (BAU), emissions to which the actual GHG emissions have to be compared. When employing the synthetic control method (SCM), the counterfactual for each "treated" country (i.e., Annex B country that ratified the KP and, thus, is subject to GHG emission targets) is constructed by a weighted average of "non-treated" regions (i.e., regions without binding emission targets under the KP) such that the actual country and its synthetic counterpart coincide as much as possible with respect to emissions before the "treatment" and in all relevant economic characteristics that are unaffected by it. The difference of the emission paths of the actual country and its synthetic counterpart following the treatment reveals the influence the binding emission targets of the KP imposed on the development of GHG emissions. The identification of the "true" counterfactual hinges crucially on the availability of appropriate controls in the control group (also called donor pool). Given that Annex B countries differ considerably from non-Annex B countries in many important characteristics, we run an additional specification using US state level data in the donor pool.

We find no statistically significant and persistent treatment effect for any of the Annex B countries under investigation. This holds no matter whether we consider for the time of treatment (i) the year of ratification of each Annex B country (e.g., 2002 for the EU countries), (ii) the adoption of the Kyoto Protocol in 1997 or (iii) the year 2005 when the Protocol entered into force. We provide evidence, however, that more drastic interventions like the collapse of the former Soviet Union did have a significant effect on the CO_2 emissions of Poland (see Section 3.4). We also find that – overall – the use of US states as the control group is preferable, as the emission paths are very similar to those of Annex B countries. Finally, we argue that the applied econometrician faces very similar challenges when analyzing other types of international environmental policies. As a consequence, the application of the SCM, which allows us to address these challenges adequately, may be preferable to more traditional panel data approaches that are likely to produce invalid results.

2 Greenhouse Gas Emission Targets under the Kyoto Protocol

In the Kyoto Protocol (KP), initially adopted on 11 December 1997, 37 industrialized countries (and the European Union), so called Annex B countries, committed to reduce the emissions of six greenhouse gases (GHGs) by 5.2% on average over the period between 2008 to 2012 compared to 1990 levels. More specifically, each of the 37 Annex B countries accepted individual emission targets that had to be met by the end of the first commitment period.

In this paper we aim to assess the effectiveness of the KP with respect to its primary target, namely the reduction of GHG emissions in Annex B countries with binding emission targets. Therefore, we consider the ratification of the KP by a country with a binding emission target under the KP as a "treatment" and investigate the effect of this treatment. To answer this question it is crucial to identify the counterfactual GHG emissions that each treated country would have had emitted were they not subject to the treatment and compare these to the actual GHG emissions of the respective country. The resulting difference, the treatment effect, is a measure of the effectiveness of the KP in the sense that it elicits the efforts of a particular country to reduce GHG emissions given that the KP entered into force.[3]

To elicit this treatment effect we are confronted with two major obstacles. First, while it is clear which countries we consider as treated, i.e. countries with binding emission targets under the KP that also ratified it, the exact timing of this treatment for each country is much less obvious. One might argue that the date of ratification is the date of treatment, as only from that point onwards a country adopted the emission target into national legislation and, thereby, certifies that it considers the emission target imposed by the KP as binding. In fact, we shall use the date of ratification as the time of treatment in our main specification.[4] However, there are two other plausible dates for the time of treatment:

1. We consider December 1997, when the KP was adopted, as the earliest time that

[3]Note that this treatment effect cannot elicit any effect that the KP (or its enaction) might have had on the emissions of all countries (e.g., due to a increase in the public awareness of the matter), treated or non-treated. In this sense the treatment effect may underestimate the effectiveness of the KP. To estimate this "total" effect, we would need to observe a twin Earth that is identical to our Earth in all aspects short of the enaction of the KP. Obviously, there is no way to construct such an ideal counterfactual.

[4]Throughout the paper, we consider the year of the treatment event as the treatment year if the event took place in the third or forth quarter and we consider the year before the event as the treatment year if the event took place in the first or second quarter.

these emission reduction targets may already have influenced the countries that finally ratified the KP, as the emission targets were already known then. As a consequence, we run an additional specification with 1997 as the time of treatment.

2. After the US withdrew its support in 2001 it was not clear whether the KP will actually enter into force, because of the two hurdles the KP had to take.[5] Due to this uncertainty, even countries that ratified the KP may have been reluctant to take costly measures to reduce GHG emissions before the KP entered into force. With Russia's ratification in November 2004 the protocol had taken both hurdles and became effective in early 2005. As a consequence, we run an additional specification where we consider the year 2004 as the time of treatment.

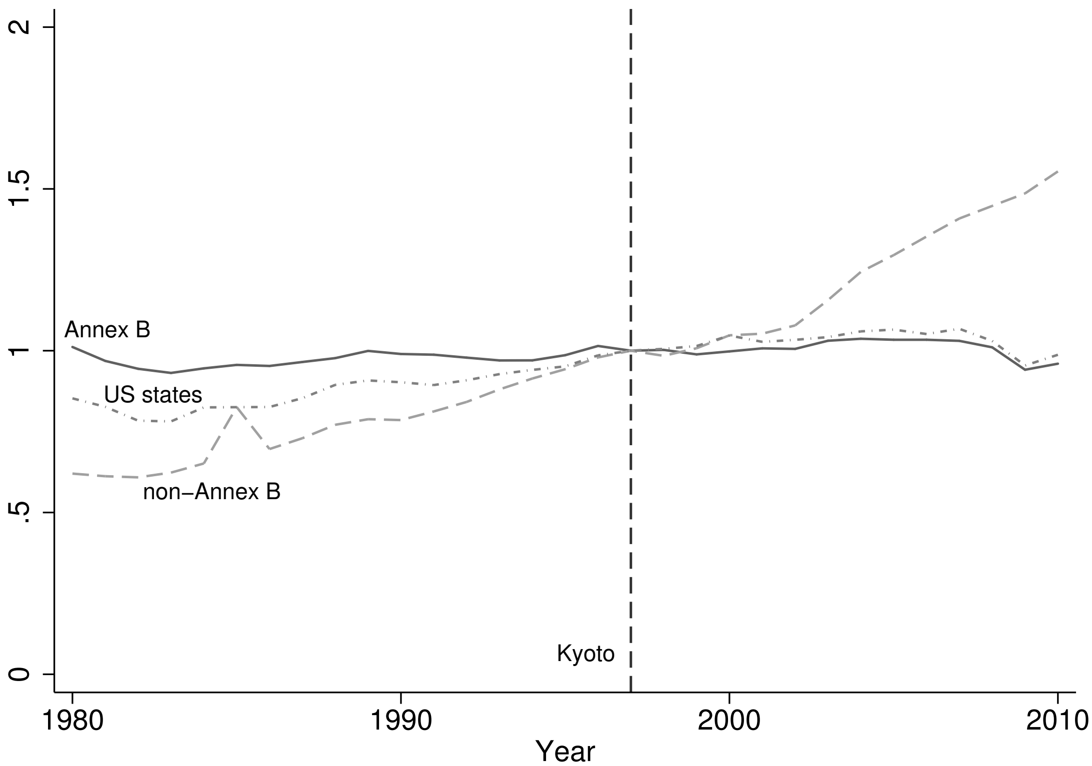
Second, to elicit the treatment effect it is crucial that there are other countries (or, more general, regional entities) that did not receive the treatment. In our case these can be regions without any binding GHG emission targets under the KP or countries that would have had binding emission targets under the KP but did not ratify it. If all countries (regional entities) were equal in all respects apart from receiving the treatment or not, the treatment effect would simply be given by the difference in GHG emissions of treated and non-treated countries. Of course, not all countries are alike. Even worse, there is a clear selection bias with respect to the treatment: Annex B countries roughly cover the industrialized world.

Figure 1 shows the aggregated (average) CO_2 emissions of Annex B and non-Annex B countries relative to their 1997 emissions between 1980 and 2010. We observe that CO_2 emissions were relatively stable for Annex B countries, while they increased considerably for non-Annex B countries in particular after 2000. While the different development in the post-treatment period, i.e. after 1997, may be the effect of the treatment, the differences in the pre-treatment period give rise to serious concerns. The key identifying assumption in settings where we observe the outcome of interest for treated units and controls prior and after the intervention (including difference-in-differences) is the common trend assumption. It says that – in the absence of the treatment – treated units and controls must share a common trend.[6] Although it is impossible to directly test for a common trend without imposing strong assumptions about the treatment effect, the drastic and statistically significant differences in pre-treatment trends shown in Figure 1 clearly indicate a

[5] According to the rules of the KP, it only enters into force if at least 55 countries ratify it and ratifying countries account for at least 55% of 1990 GHG emissions from Annex B countries.

[6] Differences in the absolute value of the outcome of interest and (to some extend) covariates are less of a problem, as they can be absorbed by the use of fixed-effects.

Figure 1: Development of average CO_2 emissions for Annex B countries, non-Annex B countries and US states



Note: Graph shows aggregated CO_2 emissions relative to 1997 emissions for all Annex B countries, non-Annex B countries for which data is available for the period 1980–2010 and US states. For Annex B countries this excludes many Eastern European countries (e.g., Russia and the Ukraine), as data start in 1990 or later. Including Eastern European countries would lead to a significant dip starting in the late 1980s/early 1990s due to the collapse of the Former Soviet Union, as shown in Grunewald and Martinez-Zarzoso (2015).

violation in case of the KP. Apart from the differences in trends, non-Annex B countries also differ significantly with respect to other relevant country characteristics, as we shall show in Section 3.2.

Non-Annex B countries face an additional problem that may invalidate them as suitable controls. According to the rules of the KP, Annex B countries may achieve their emission targets either via domestic emission reductions or through the use of one of the three flexibility mechanisms: emissions trading, joint implementation and clean development mechanism. Although the KP clearly states that the use of flexibility mechanisms should only be supplemental to domestic emission reduction efforts, they may blur the distinction between treated and non-treated countries. In particular the clean development mechanism encourages Annex B countries to cooperate with non-Annex B countries to reduce GHG emissions of non-Annex B countries. These emission reductions in non-Annex B countries can then be credited to the reduction target of the supporting Annex B countries. Thus, also the emissions of non-Annex B countries may be influenced by the KP, even though they do not face any direct reduction obligations.

In order to circumvent the problems associated with using non-Annex B countries as the control group to estimate counterfactual GHG emissions for treated countries, we also use US state level data in an alternative specification. US states are untreated, as the US did not ratify and, thus, never had any binding obligations under the KP. US state level data comes with several key advantages: First, the differences in trends between CO_2 emissions of US states and Annex B countries are rather moderate, as shown in Figure 1. Second, as US states clearly belong to the industrialized world, the differences in other key country characteristics are more moderate, although still significant (see Section 3.2). Finally, as the US did not ratify the KP, they are also not part of any of the flexibility mechanisms.

3 Empirical Analysis

In the following, we aim at eliciting the treatment effect of being committed to a specific emission target under the KP for each treated country individually. The main challenge for estimating such an effect is a missing data problem (Rubin, 1976), as we cannot observe a particular country having both a binding emission target and no emission target at the same time. Moreover, countries with and without binding emission targets may differ systematically with respect to both their emission paths and other important country characteristics. Therefore, average CO_2 emissions of countries with targets cannot be simply compared to average emissions of countries that have none, i.e. the assignment of emission targets cannot be treated as random (Rubin, 1976, 1978, 2005).

3.1 The Synthetic Control Method

There are several potential strategies to solve this problem (Imbens and Wooldridge, 2009). In cases as ours, where different groups are either exposed or not exposed to some kind of treatment over a certain time period, the most often applied method is a differences-in-differences (DiD) approach (Bertrand et al., 2004) or an extension of it. We employ the synthetic control method (SCM) developed by Abadie and Gardeazabal (2003), Abadie et al. (2010), and Abadie et al. (2015). It can be seen as an extension of a standard DiD and allows us to account for the various empirical challenges described above. More specifically, the synthetic control approach exhibits three key advantages over classical DiD estimation that renders it particularly suitable for the present research question.

First, SCM is able to re-balance treated units and controls with respect to both (i) the pre-treatment development of the CO_2 emissions themselves and (ii) important predictors of selection into the treatment, i.e. being part of the industrialized world. Second, SCM allows to estimate the counterfactual emissions path for every single country and every year following the adoption of the KP. Thus, we do not only get an average effect for all countries and all time periods under investigation, but we are also able to identify country-year-specific developments and characteristics. We are therefore able to identify potentially heterogeneous treatment effects which may be, for example, the result of highly heterogeneous targets. Third, and most important, the standard DiD approach faces an additional challenge when analyzing counterfactual outcomes. On the one hand, a country's level of GHG emissions depends on several socio-economic factors that should be controlled for in a regression analysis to avoid an omitted variable bias. On the other hand, almost all of these factors may also be influenced by the treatment and therefore constitute bad controls according to the definition of Angrist and Pischke (2008).[7] In contrast to DiD, SCM is flexible in the sense that one can control for these factors using only pre-treatment information. As a consequence, the resulting counterfactual will account for important pre-treatment socio-economic characteristics of countries' which are independent of the treatment.

Suppose that there are $J+1$ countries where $j = 1$ denotes the treated country – which, in our case, corresponds to a binding emission target under the KP – and $j = 2, \ldots, J+1$ are all untreated countries or US states in the donor pool. In addition, let T_0 be the time of treatment. For the treated country we have data about the actual emission path (Y_{1t}), but we are ignorant about the counterfactual emissions which would have occurred if this country would not have been subject to the treatment (Y_{1t}^N for $t > T_0$). Thus, we have to find an estimate for Y_{1t}^N to obtain an estimate for the treatment effect α_{1t}:

$$\alpha_{1t} = Y_{1t} - Y_{1t}^N \ . \tag{1}$$

Abadie and Gardeazabal (2003), Abadie et al. (2010), and Abadie et al. (2015) propose to make use of the observed characteristics of the countries in the control group or donor pool.

[7] For example, investments to reduce GHG emissions may have an impact on GDP per capita, CO_2 intensity, electricity production, etc.

The underlying idea is to find weights $W = (\omega_2, ..., \omega_{J+1})'$, with $\omega_j \geq 0$ for $j = 2, \ldots, J+1$ and $\sum_{j=2}^{J+1} \omega_j = 1$, such that the weighted average of all countries in the donor pool resembles the treated country with respect to GHG emissions in the pre-intervention period and all other relevant aspects (Z).

Formally, we seek W such that:

$$\sum_{j=2}^{J+1} \omega_j^\star Y_{jt} = Y_{1t} \quad \text{for all } t < T_0 \qquad \text{and} \qquad \sum_{j=2}^{J+1} \omega_j^\star Z_j = Z_1 \ . \tag{2}$$

Then $\sum_{j=2}^{J+1} \omega_j^\star Y_{jt}$ for $t \geq T_0$ is an estimate for the unobserved counterfactual emissions path Y_{1t}^N inducing an estimate for the treatment effect:

$$\widehat{\alpha}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} \omega_j^\star Y_{jt} \ , \quad t \geq T_0 \ . \tag{3}$$

In general, a vector W such that equations (2) hold may not exist (in particular, if there are structural differences between treated countries and controls). However, one can choose the weights such as to

$$\min_W (X_1 - X_0 W)' V (X_1 - X_0 W) \ , \tag{4}$$

where X_1 denotes a $(k \times 1)$ vector of pre-intervention characteristics of the treated country, which may include the pre-intervention emission path, and X_0 denotes a $(k \times J)$ matrix of the same variables for the J countries in the donor pool (Abadie and Gardeazabal, 2003; Abadie et al., 2010, 2015). The symmetric and positive definite matrix V weights the relative importance of the various characteristics included in X. Obviously, the optimal weights W depend on the weighting matrix V. We follow Abadie et al. (2010) in choosing V by using a regression based method and equal weights. For further discussion on the synthetic control method including several extensions, see Abadie et al. (2010) and Abadie et al. (2015).

In other words, we use SCM to create the specific counterfactual country for each treated country via a convex combination of all units in the donor pool. To increase the comparability of countries we normalize the outcome of interest (CO_2 emissions) with the year of treatment T_0 as the base year (Cavallo et al., 2013).

Our set of predictors when using non-Annex B countries in the donor pool includes: Several years of normalized (treatment year=1) CO_2 emissions[8]; the averages of two 5-year

[8]More specifically, we use 1981, 1983, 1985, ...

periods prior to the treatment for all predictors (e.g., 1987-1991 and 1992-1996 for 1997 as treatment), i.e. GDP per capita, GDP growth, human capital index, life expectancy, agricultural, industry and services value added and population growth; and the averages for two sub-periods of the post-treatment periods (e.g., 2000 - 2005 and 2006-2010) for life expectancy, human capital index and population growth. In doing so, we assume that the latter three variables will not be affected by the treatment and, therefore, do not constitute bad controls in the sense of Angrist and Pischke (2008).

Unfortunately, several of the above listed variables are not available for US states. Thus, we rely on several years of normalized (treatment year=1) CO_2 emissions, GDP per capita, GDP growth, and population growth when using US states as the donor pool. As we do expect differences in the classification and coding between country level and US state level data, in particular for GDP, we use the changes rather then the levels (as for non-Annex B countries) with respect to the treatment year values.[9] More specifically, we use normalized per capita GDP, GDP growth and population growth for, e.g., 1987–1991 and 1992–1996 and, in addition, population growth data for 2000–2005 and 2006–2010 (1997 as treatment year).

As the SCM does not provide classical standard errors to infer statistical significance, Abadie and Gardeazabal (2003), Abadie et al. (2010), and Abadie et al. (2015) suggest to run placebo or permutation tests. The underlying idea is to estimate counterfactual emission paths for regional entities in the donor pool, i.e., for regions without any treatment. In an ideal world with the perfect analogue of the "treated" country being available in the donor pool, we would find no treatment effects for all countries in the donor pool and all post-treatment years, as the countries in the donor pool, i.e. the control group, did not receive any treatment. However, in practice we will always find placebo treatment effects to at least some extend. As a consequence, we only consider the actual treatment effect to be statistically significant if it is significantly larger than the placebo treatment effects.

Our inference approach rests on a combination of two refinements of the classical placebos studies that have been proposed by Abadie et al. (2010) and Abadie et al. (2015). First, we run a placebo treatment on all countries in the donor pool and select the top 19 countries/US states in terms of pre-treatment root mean squared prediction error (RMSPE) to avoid rejecting the significance of treatment effects on the basis of outliers within the placebos studies. Second, we calculate the ratio between the treatment effect (the root mean squared treatment effect) and the pre-treatment root mean squared

[9]We expect these differences to be constant over time and therefore focus on the changes in order to eliminate any inconsistencies.

prediction error (RMSPE) for all 20 countries (the top 19 countries from the placebo treatment plus the treated country) for each year of the post-treatment period. For each year of the post-treatment period, we then calculate a probability that resembles the relative frequency that a randomly chosen country out of the 20 countries in the placebo test has a RMSPE ratio that is at least as large than that of the treated country. We use a combination of the two refinements as proposed by Abadie et al. (2010) and Abadie et al. (2015), as placebos with large pre-treatment RMSPE, i.e. a poor fit, are likely to experience large placebo treatment effects in the post-treatment period and, therefore, lead not only to (i) large treatment/placebo effects but also (ii) large RMSPE ratios. Finally, we consider the treatment effect of the treated country for a particular post-treatment year to be significant if none of the 19 countries with placebo treatments shows a larger RMSPE ratio than the actual treated unit, i.e. the relative frequency of finding such an effect by randomly drawing one out of the 20 countries is $1/20 = 0.05$.

3.2 Data

We analyze the effect of being committed to an emission target under the KP for 15 major Western GHG emitters. Eastern European countries are excluded from the empirical analysis due to data availability.[10] The 15 treated countries under investigation are: Australia, Austria, Belgium, Canada, Finland, France, Germany, Great Britain, Italy, Japan, Netherlands, Norway, Portugal, Spain, and Sweden. With respect to GHG emissions, these countries are responsible for approximately 50% of total 1990 GHG emissions of the countries with binding emission targets under the KP. In order to reduce the imbalance in country characteristics between Annex B and non-Annex B countries as much as possible, we restrict the country level control group to countries being classified as high income and upper middle income countries by the World Bank.

The data used in the empirical analysis stem from several different sources. Data on country-level CO_2 emissions, value added for agriculture, industry and services, GDP growth, life expectancy and population growth are taken from the World Development Indicators published by the World Bank. GDP per capita and the human capital index originate from the Penn World Tables published by the University of Groningen. Additional information on the KP (list of Annex B countries with targets) stems from the

[10]Moreover, at the time of adoption of the KP in 1997, Eastern European countries exhibited emission levels far below their emission targets due to the severe economic downturn during the 1990s which followed the breakdown of the Former Soviet Union (FSU). As these countries were not expected to reach emission levels at or even above their Kyoto targets in the near future – despite their economic recovery –, they had little economic incentives to reduce emissions.

Table 1: Descriptive Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N	P-values
Annex B countries (under investigation)						
GDP per capita	23439.94	5485.69	10401.74	43569.25	165	
GDP growth	2.15	2.09	-6.51	7.60	165	
Human Capital Index	2.77	0.29	2.19	3.3	165	
Life Expectancy	76.85	1.31	73.67	80.42	165	
Agri. value added*	3.52	1.82	1.22	13.93	165	
Ind. Value added*	31.05	3.1	23.56	38.41	165	
Serv. valued added*	65.42	3.21	58.46	73.23	165	
Population growth	0.48	0.39	-0.22	1.78	165	
High income and upper middle income Non-Annex B countries^{+}						
GDP per capita	10857.28	10709.68	335.49	67883.89	694	0.00
GDP growth	2.13	7.50	-42.62	90.88	813	0.07
Human Capital Index	2.36	0.37	1.68	3.52	511	0.00
Life Expectancy	69.83	5.99	40.97	80.13	823	0.00
Agri. value added*	11.24	10.21	0.00	67.38	687	0.00
Ind. Value added*	33.20	12.64	8.02	74.67	678	0.13
Serv. valued added*	55.49	14.21	8.24	90.63	679	0.00
Population growth	1.81	1.70	-6.49	16.51	994	0.00
US states						
GDP per capita	27514.78	9703.95	15468	87544	561	0.03
GDP growth	1.98	2.57	-12.90	11.90	561	0.00
Population growth	0.01	0.01	-0.04	0.07	561	0.00

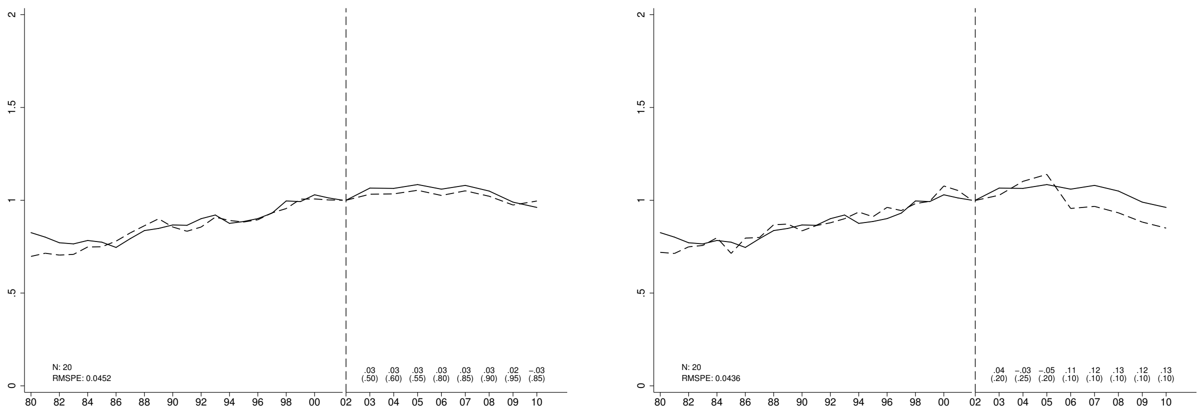
Note: Descriptive statistics for each (sub-)sample and the 11 years (1987–1997) prior to the treatment, i.e. the adoption of the KP. The p-values in the last column are based on t-tests for the equality of means between Annex B countries and the two potential control groups.
*: in % of GDP, $^{+}$: not sufficient data available for some non-Annex countries, i.e. they are not part of the donor pool.

United Nations Framework Convention on Climate Change (UNFCCC). US state level information on CO_2 emissions is taken from the US Energy Information Administration (EIA), real GDP per capita and GDP growth stem from the Bureau of Economic Analysis (BEA) and population data is based on the Population Estimates Program (PEP) of the United States Census Bureau. Unfortunately, data on the remaining predictors (human capital index, life expectancy, etc.) are not readily available at the US state level. In order to avoid any problems caused by potential inconsistencies in the classification of country and US state level data, we only use changes over time for the available predictors when using US states as a control group in the later analysis. Table 1 shows the summary statistics for all data used in the empirical analysis.[11]

[11] Although CO_2 data for US states is available until 2012, there is still no consistent information for CO_2 emissions of Annex B countries after 2010. We are therefore – in principle – able to estimate the counterfactual emission path until 2012, i.e. until the end of the first commitment period but do not have

Figure 2: Synthetic Control for Canada based on

a) non-Annex B countries b) US states



Note: Actual and synthetic CO_2 emission path for Canada. The solid line stands for the actual Canadian CO_2 emissions whereas the dashed line represents the synthetic Canada, i.e. the counterfactual emission path, based on non-Annex B country data (left) and US state level data (right). N is the number of countries in the donor pool used for the placebo studies and we display the figures for the pre-treatment RMSPE. The numbers underneath the lines indicate the treatment effects in percent and the corresponding probabilities in parenthesis (see section 3.1).

Table 1 clearly shows the predictor imbalance for Annex B, non-Annex B countries and US states for the 11 years prior to the adoption of the KP, i.e. 1987–1997. As discussed earlier, it is evident that there is a statistically significant difference for most predictors (except industry value added) between the two groups of countries. Moreover, there is also a significant difference between Annex B countries and US states. However, in contrast to non-Annex B countries per capita GDP is actually higher and population growth lower (on average) for US states than for Annex B countries. GDP growth, however, is very similar for all groups in the period of 1987–1997.

3.3 Results

In the following, we report the results for the two different specifications, where either selected non-Annex B countries or US states act as non-treated control entities from which we construct the counterfactual CO_2 emission paths of the 15 Annex B countries under investigation. We consider ratification as the treatment event.

As a representative example, Figure 2 shows the normalized CO_2 emissions path of Canada (solid line) and its synthetic counterfactuals (dashed line) based on non-Annex B countries (left) and US states (right). We observe that for both specifications the synthetic

information on the actual emissions of Annex B countries to estimate the treatment effect.

controls for Canada matches actual CO_2 emissions very well in the pre-treatment period (up to 2002). This is also evident from the small RMSPE values of 0.0452 and 0.0436. Also the general pattern of counterfactual emissions in the post treatment period (after 2002), i.e. the CO_2 emissions Canada had if it were not subject to a binding emission target under the KP, is similar for both specifications. Actual CO_2 are slightly above the counterfactual emissions during the post treatment period. However, we do observe slight differences between the two specifications. US state level data predicts lower counterfactual CO_2 emissions of Canada in the post treatment period. In fact, the treatment effect ranges from -3% to +3% for non-Annex B countries and -5% to +13% for US state level data. Note that a positive (negative) number indicates actual emissions are higher (lower) than the emissions of the synthetic control. Thus, a treatment effect in the sense that countries with a binding emission target under the KP experience lower actual emissions than predicted by their counterfactuals would result in negative numbers.

As an indicator of significance, the values in parenthesis give the probabilities that if one randomly draws one country out of the 20 countries in the placebo study (the top 19 non treated countries with respect to low RMSPE in the pre-treatment period plus the treated country) one would draw a country that exhibits at least as high a deviation from their actual normalized CO_2 emissions than the treated country. We do not find any significant effects both when using non-Annex B country level data and US state level data to construct counterfactual emissions. Thus, we do not find any evidence that the adoption of binding emission targets under the KP had any significant CO_2 emissions reduction effect for Canada.

Tables 2 and 3 summarize the results for all 15 Annex B countries under investigation and for the two model specifications (using non-Annex B countries and US states in the donor pool). The numbers indicate the deviation of the actual normalized CO_2 emissions from its synthetic counterfactuals, i.e. the numbers for Canada are identical to the numbers shown in Figure 2. Likewise, numbers in parenthesis show probabilities as discussed in section 3.1 and, thus, reflect the statistical significance of the effect.[12]

For the specification using non-Annex B countries to construct counterfactual CO_2 emissions (Table 2), we find few negative treatment effects being mostly in single digits with the biggest effect for Portugal (-17%) in the years 2006, 2008 and 2009. None of the negative effects are statistically significant (corresponding to a probability of 0.05). However, we do find quite substantial positive effects for several countries including Austria,

[12]Analogous graphs to Figure 2 for all countries and both specifications and the placebo studies to infer statistical significance are in the Appendix (Figures 5–8).

Table 2: Estimates for Treatment Effects based on non-Annex B Countries

Country	Year 2002	2003	2004	2005	2006	2007	2008	2009	2010
Australia							.05	.12	.02
							(.38)	(.69)	(.92)
	RMSPE: 0.0741								
Austria	.06	.27	.24	.23	.25	.36	.39	.22	.53
	(.45)	(.10)	(.25)	(.25)	(.25)	(.20)	(.25)	(.55)	(.30)
	RMSPE: 0.0858								
Belgium	-.05	.13	.09	.04	.03	.11	.08	.08	.26
	(.50)	(.35)	(.50)	(.90)	(.85)	(.80)	(.75)	(.70)	(.55)
	RMSPE: 0.1210								
Canada		.03	.03	.03	.03	.03	.03	.02	-.03
		(.50)	(.60)	(.55)	(.80)	(.85)	(.90)	(.95)	(.85)
	RMSPE: 0.0452								
Finland	.11	.33	.29	.01	.26	.34	.18	.10	.41
	(.15)	(.05)	(.10)	(1.0)	(.10)	(.20)	(.55)	(.70)	(.30)
	RMSPE: 0.1032								
France	-.01	.21	.27	.23	.23	.46	.42	.29	.55
	(.90)	(.10)	(.25)	(.30)	(.35)	(.20)	(.25)	(.45)	(.30)
	RMSPE: 0.1607								
Germany	-.04	.04	.03	-.01	.02	.05	.06	.02	.06
	(.60)	(.75)	(.80)	(1.0)	(.90)	(.90)	(.85)	(.90)	(.75)
	RMSPE: 0.1443								
Great Britain	-.01	.20	.21	.16	.20	.45	.45	.25	.54
	(.90)	(.15)	(.40)	(.50)	(.40)	(.20)	(.25)	(.55)	(.30)
	RMSPE: 0.1130								
Italy	.05	.20	.22	.16	.21	.38	.39	.18	.39
	(.60)	(.10)	(.35)	(.50)	(.35)	(.20)	(.25)	(.60)	(.40)
	RMSPE: 0.0935								
Japan	.04	.09	.12	.07	.09	.14	.13	.05	.15
	(.60)	(.45)	(.45)	(.70)	(.65)	(.60)	(.60)	(.80)	(.60)
	RMSPE: 0.0664								
Netherlands	.04	.15	.18	.12	.11	.24	.27	.24	.43
	(.60)	(.20)	(.40)	(.50)	(.55)	(.45)	(.40)	(.55)	(.30)
	RMSPE: 0.0885								
Norway	-.08	.06	.04	.02	.09	.06	.21	.20	.46
	(.30)	(.70)	(.75)	(1.0)	(.60)	(.90)	(.40)	(.55)	(.30)
	RMSPE: 0.0952								
Portugal	.11	.00	-.07	-.08	-.17	-.10	-.17	-.17	-.15
	(.15)	(.95)	(.50)	(.50)	(.30)	(.70)	(.40)	(.55)	(.55)
	RMSPE: 0.0461								
Spain	.08	.03	.09	.10	.11	.13	.05	-.05	-.08
	(.30)	(.75)	(.45)	(.50)	(.50)	(.50)	(.85)	(.80)	(.75)
	RMSPE: 0.0916								
Sweden	.11	.13	.14	.06	.04	.08	.11	.00	.26
	(.15)	(.20)	(.40)	(.80)	(.85)	(.90)	(.65)	(1.0)	(.55)
	RMSPE: 0.1931								

Note: The Table contains treatment effect estimates for each Annex B country under investigation using non-Annex B countries to construct the synthetic counterfactuals and considering the year of ratification as the time of treatment. For each country, we display the yearly treatment effects in percent and the respective probabilities of finding such an effect in parenthesis. Finally, we report the (pre-treatment) RMSPE for each country.

Table 3: Estimates for Treatment Effects based on US States

Country	Year 2002	2003	2004	2005	2006	2007	2008	2009	2010
Australia							.05	.12	.10
							(.15)	(.10)	(.10)
	RMSPE: 0.0769								
Austria	.07	.10	.06	.05	.14	.07	.11	.20	.33
	(.15)	(.05)	(.15)	(.35)	(.10)	(.30)	(.20)	(.10)	(.05)
	RMSPE: 0.0513								
Belgium	-.03	-.00	-.06	-.12	-.03	-.10	-.04	.19	.29
	(.40)	(.85)	(.15)	(.05)	(.35)	(.20)	(.55)	(.10)	(.05)
	RMSPE: 0.1063								
Canada		.04	-.03	-.05	.11	.12	.13	.12	.13
		(.20)	(.25)	(.20)	(.10)	(.10)	(.10)	(.10)	(.10)
	RMSPE: 0.0436								
Finland	.12	.22	.15	-.10	.22	.15	.06	.21	.45
	(.05)	(.05)	(.05)	(.10)	(.05)	(.05)	(.40)	(.10)	(.05)
	RMSPE: 0.0881								
France	.01	-.01	-.00	-.04	.02	-.04	.00	.28	.33
	(.90)	(.70)	(.85)	(.40)	(.55)	(.55)	(.95)	(.05)	(.05)
	RMSPE: 0.1164								
Germany	-.05	.01	-.02	-.03	.15	.08	.17	.09	.07
	(.25)	(.80)	(.45)	(.60)	(.10)	(.35)	(.10)	(.25)	(.45)
	RMSPE: 0.0454								
Great Britain	-.01	-.03	-.04	-.08	.02	-.05	-.01	.17	.25
	(.90)	(.45)	(.25)	(.15)	(.60)	(.50)	(.95)	(.10)	(.05)
	RMSPE: 0.0586								
Italy	.05	.04	.01	-.03	.10	.05	.06	.12	.18
	(.25)	(.30)	(.80)	(.55)	(.20)	(.50)	(.45)	(.20)	(.10)
	RMSPE: 0.0454								
Japan	.01	.04	.04	.03	.03	.01	-.02	-.08	-.04
	(.85)	(.40)	(.25)	(.60)	(.40)	(.85)	(.65)	(.25)	(.60)
	RMSPE: 0.0214								
Netherlands	.05	.04	.04	-.01	.03	.03	.10	.29	.39
	(.30)	(.40)	(.25)	(.70)	(.35)	(.65)	(.20)	(.05)	(.05)
	RMSPE: 0.0598								
Norway	-.04	.05	-.01	-.06	.14	.14	.32	.39	.77
	(.35)	(.25)	(.80)	(.20)	(.10)	(.05)	(.05)	(.05)	(.05)
	RMSPE: 0.1022								
Portugal	.15	-.00	-.06	-.07	.02	.04	.02	.04	.00
	(.05)	(.95)	(.15)	(.20)	(.70)	(.60)	(.70)	(.65)	(1.0)
	RMSPE: 0.0803								
Spain	.13	.10	.07	.07	.26	.29	.21	.12	.11
	(.05)	(.05)	(.10)	(.20)	(.05)	(.05)	(.05)	(.15)	(.35)
	RMSPE: 0.0709								
Sweden	.12	.08	.07	-.01	.06	-.02	.08	.17	.41
	(.05)	(.10)	(.15)	(.80)	(.25)	(.80)	(.40)	(.10)	(.05)
	RMSPE: 0.1120								

Note: The Table contains treatment effect estimates for each Annex B country under investigation using US states to construct the synthetic counterfactuals and considering the year of ratification as the time of treatment. For each country, we display the yearly treatment effects in percent and the respective probabilities of finding such an effect in parenthesis. Finally, we report the (pre-treatment) RMSPE for each country.

Finland, France, Great Britain, Italy, The Netherlands and Norway. At the same time, we only find a significant positive treatment effect (+33%) for Finland in the year 2003.

When looking at the specification using US state level data to construct counterfactual CO_2 emissions (Table 3), the general pattern is that effects tend to be smaller in magnitude but are more often statistically significant. As a result, we do now find a statistically significant negative effect for Belgium in 2005 (-12%). In line with our findings based on non-Annex B countries, however, we find more pronounced positive effects. There are positive and significant effects for Austria (2 years), Belgium (1 year), Finland (6 years), France (2 years), Great Britain (1 year), The Netherlands (2 years), Norway (4 years), Portugal (1 year), Spain (5 years) and Sweden (2 years).[13]

In summary, we do not find any evidence that a binding emission target under the KP induced a significant and persistent emission *reduction* effect for Annex B countries with binding emission targets under the KP, although we do find significant negative effects for one country in one year (of the nine year post treatment period) in one of our specifications.[14] This finding is further illustrated in Figure 3 showing the average normalized CO_2 emissions of the Annex B countries under investigation (solid line) and the average counterfactual CO_2 emissions path (dashed line) based on non-Annex B countries (left) and US states (right). We observe that the average counterfactual CO_2 emissions path fits the average actual emissions path quite well not only in the pre-treatment but also in the post treatment period. If at all, we find that actual emissions are higher (not lower!) than the counterfactual GHG emissions, i.e. the synthetic controls based on comparable countries and US states from the control group.

When comparing Figures 3 and 1 it is also evident that the synthetic control method successfully balanced the pre-treatment trends, i.e. the opposing trends between Annex B and non-Annex B countries, as shown in Figure 1, vanished entirely in Figure 3.

3.4 Robustness Checks

For all countries under investigation we consider the ratification of the KP as the time of treatment. However, as the emission targets were already known in 1997, it might be that the adoption of the KP already induced changes in the emission paths of treated

[13]Given the importance of Delaware and Nevada for many treated countries (see Table 5 in the Appendix), we estimated an alternative specification excluding both states from the donor pool. Again, the results look very much the same as the baseline specification above with a significant negative effect only for Portugal in 2010, which also shows a considerably larger RMSPE.

[14]However, we do find significant and persistent positive treatment effects for Finland, Norway and Spain (positive treatment effects in four or more years out of the 9 year post treatment period).

Figure 3: Average actual and synthetic CO_2 emissions for synthetic controls based on

a) non-Annex B countries b) US states

Note: The solid line represents the (average) normalized (treatment year=1) CO_2 emission of the Annex B countries under investigation. The dashed line line indicates the average normalized emissions of the synthetic controls for the treated countries based on non-Annex B countries (left) and US states (right).

countries. In fact, we consider 1997 as the earliest time at which the KP could have imposed a treatment effect. As a consequence, we run an additional specification, where we assume 1997 as the year of treatment. This robustness check could also be interpreted as a placebo in time analysis (Abadie and Gardeazabal, 2003; Abadie et al., 2010, 2015). Results are shown in Tables 6 and 7, and Figures 9 and 10 in the Appendix and are very similar to our main specification. We do not find any significant treatment effects using non-Annex B countries in the donor pool and hardly find any significant negative treatment effects when using states in the donor pool (Finland and The Netherlands in the year 2000 being the only exception). Again, we do find some positive treatment effects, in particular for Norway and Spain.

Although an Annex B country might have known its emission reduction target under the KP already in 1997, a country might not have taken any action to reduce its GHG emissions prior to the date the KP entered into force, even if the country has ratified the KP before. Of course, this does not invalidate the analysis when choosing ratification or even adoption as the time of treatment. If the synthetic counterfactual matches the country under consideration well, all we should see is that counterfactual emissions start to deviate from actual emissions not at the considered time of treatment but at some later time at which emission reduction efforts started. Yet, taking too early a treatment date comes at the disadvantage that fewer pre-treatment periods of data are available to construct the synthetic counterfactuals which might result in a poorer match between actual country and synthetic counterfactual. As a consequence, we run an additional specification were we consider 2004 as the time of treatment, because subsequent to Russia's ratification in November 2004 the KP entered into force in February 2005. Results are shown in Table 8, and Figures 11 and 12 in the Appendix and are highly consistent with the results of our two main specifications.

For the 15 Annex B countries investigated we do find very little evidence for a persistent treatment effect, i.e. a permanent reduction in CO_2 emissions compared to their synthetic counterfactuals no matter what time we consider as the time of treatment. One reason for this might be that CO_2 emission levels are strongly correlated with economic performance and, thus, are vulnerable to business cycle fluctuations. In other words, favorable global economic conditions for Annex B countries (at least up to 2007) could be responsible for the lack of a significant effect on emission levels although Annex B countries might have invested in cleaner production technologies. We test this hypothesis by running alternative specifications with CO_2 intensity (CO_2 emissions relative to GDP) and CO_2 emissions per capita as dependent variables, again using non-Annex B country and US state level data to construct counterfactuals for the 15 Annex B countries under consideration.[15] Tables 9–11 in the Appendix show the results of these alternative model specifications. Our findings are very consistent with our main specifications. We do not find any significant treatment effects for CO_2 intensity, and only find significant positive treatment effects in case of CO_2 emissions per capita.

We excluded Eastern European countries from the analysis for two reasons. First, we have insufficient data for most of these countries to construct proper synthetic counterfactuals. Second, after the collapse of the Soviet Union in 1989 the whole former "Eastern Block" experienced a severe economic downturn. Economic downturns are accompanied with lower production and, as a consequence, lower energy use and reduced CO_2 emissions. In fact, the collapse of the Former Soviet Union (FSU) had a significant impact on these countries' CO_2 emissions that is likely to blur any additional emission reductions due to binding emission targets under the KP.[16] To show this, we investigate the effect of the treatment "collapse of the FSU" on the CO_2 emissions of Poland.[17] Again, we construct

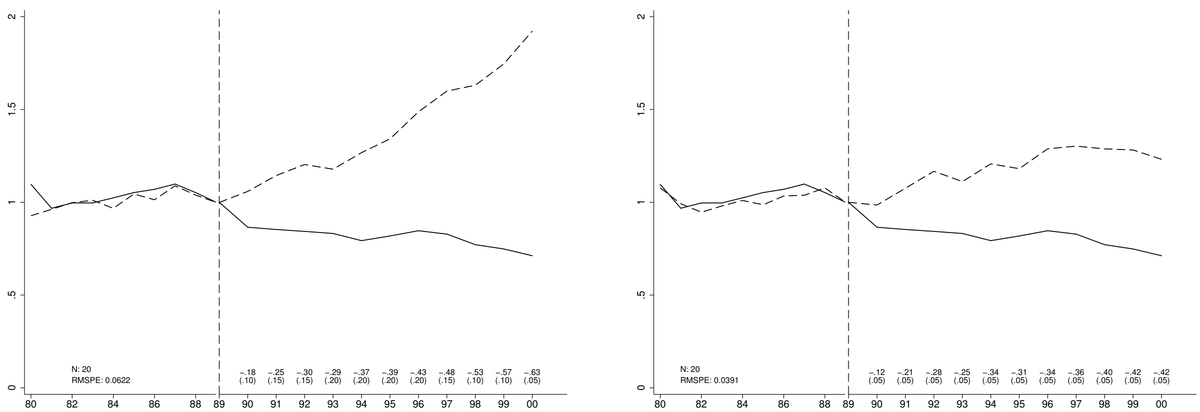
[15]Due to a discontinuity in US state level data for GDP between 1997 and 1998 we are not able to use US states in the donor pool for CO_2 intensity.

[16]Despite their economic recovery, Eastern European countries exhibited emission levels far below their emission targets at the time of adoption of the KP in 1997. As these countries were not expected to reach emission levels at or even above their Kyoto targets in the near future, they had little economic incentives to reduce emissions.

[17]Poland being one of the few Eastern European countries with sufficient data on CO_2 emissions prior

Figure 4: Poland's actual and synthetic CO_2 emissions for synthetic controls based on

a) non-Annex B countries b) US states



Note: Actual normalized (1989=1) CO_2 emissions (solid line) and synthetic control (dashed line) for Poland considering the collapse of the Soviet Union in 1989 as the treatment using non-Annex B country data (left) and US state level data to construct synthetic counterfactuals. The country weights for the synthetic control are Barbados (0.138), Fiji (0.08), Korea (0.031), Sri Lanka (0.209), Trinidad and Tobago (0.143), USA (0.193), and Vietnam (0.172) for non-Annex B countries and Illinois (0.565) and Rhode Island (0.435) for US states in the donor pool.

a synthetic counterfactual Poland (again using non-Annex B countries and US states) and compare the deviations of normalized CO_2 emissions. Figure 4 shows the results. We observe that starting from the time of the treatment in 1989 actual normalized CO_2 emissions in Poland increasingly fall short of the normalized CO_2 emissions of its synthetic counterfactual. As already seen in Section 3.3, this deviation is more pronounced but less significant using non-Annex B countries compared to US states in the donor pool. We observe a significant treatment effect at the end of the investigation period (year 2000) for synthetic controls based on non-Annex B countries and a highly significant treatment effect throughout the whole post treatment period for synthetic controls based on US states.

4 Discussion

We analyzed the effectiveness of the KP with respect to its primary goal – the reduction of domestic GHG emissions in the industrialized world – for 15 Western Annex B countries. Both in our two main specifications and in the various robustness checks we performed, we only find very little evidence for a significant and persistent emission reduction effect

to 1990.

in any of the 15 major western Annex B countries under investigation. At the same time, we do find some evidence that some Annex B countries performed even worse than comparable countries from the two donor pools. On average, the CO_2 emissions of the 15 countries are rather above than below their synthetic controls (see Figure 3). This stands in contrast to three recent empirical studies (Aichele and Felbermayr, 2012, 2013; Grunewald and Martinez-Zarzoso, 2015) that consistently find, on average, substantial (7–10%) and significant CO_2 emission reductions attributable to the adoption of binding emission targets under the KP.

Our results show the importance of addressing a number of empirical challenges when estimating the effect of international environmental policies in general, and the Kyoto Protocol in particular. The key to identifying the "true" treatment effect is the availability and the selection of appropriate controls. In order to do so, our empirical strategy differs from the previous studies in three important aspects.

First, we employ the synthetic control method (Abadie and Gardeazabal, 2003; Abadie et al., 2010, 2015). SCM enables us to construct counterfactual emissions paths for all 15 countries under investigation that reasonably match the observed emissions in the pre-treatment period (as is evident from the small root mean square prediction errors shown in Tables 2 and 3). This is possible because the treatment effects can be estimated for each country individually, which allows to individually identify the counterfactual synthetic country to the idiosyncrasies of each treated country, such as different country characteristics, emission paths, targets and ratification dates.[18]

Second, using non-Annex B country data to construct counterfactual CO_2 emission paths for the investigated Annex B countries may be problematic because of (i) irreconcilable structural differences between Annex B and non-Annex B countries including opposing pre-treatment emission paths and (ii) a bias resulting from the use of the flexibility mechanisms blurring the distinction between treated and non-treated countries. To circumvent these problems we run a second specification using US state level data to construct synthetic counterfactuals. Although US state level data comes at the disadvantage that data on some covariates is not available, it seems to be preferable to using non-Annex B country data on the ground that pre-treatment matching is considerably better (as is evident from comparing the root mean square prediction errors in Tables 2 and 3).

Third, we discard all Eastern European countries from the analysis. We consider this justified, as the former "Eastern Block" experienced a severe economic downturn after

[18]In fact, Tables 4 and 5 in the Appendix show that the countries/regions drawn from the donor pool to construct the counterfactual country differ considerably for the 15 investigated countries.

the collapse of the Former Soviet Union at the beginning of the 1990s. Even after their economic recovery, former CO_2 emission levels were not reached and, at the time of the adoption of the KP, it was evident that many Eastern European countries were subject to GHG emission targets that were unlikely to be binding (also known as “hot air”). As a consequence, these countries had no incentive to reduce GHG emissions in the first place. Nevertheless, the collapse of the FSU had a significant negative impact on the CO_2 emissions of the Eastern European countries. Incorporating these countries into the treated sample when eliciting the average treatment effect of the treated may bias the result in favor of a negative significant effect if one is not able to control for the peculiarities of those countries in particular with respect to development of CO_2 emission in the early 90s. In fact, we conjecture that considering Eastern European countries is at least partly responsible for both the size and the significance of the treatment effects reported in Aichele and Felbermayr (2012), Aichele and Felbermayr (2013) and Grunewald and Martinez-Zarzoso (2015).

Our results have strong political implications. Despite the persistent experience of failed negotiations on a successor of the KP for almost a decade, the international community seems to insist on a “Kyoto Protocol XXL”, i.e. a treaty similar to the KP but also including binding emission reduction targets for some of the former non-Annex B countries, in particular, countries in transition such as China, India, Brazil etc. Recent empirical evidence seems to support such an approach, as these studies find that – at least in average – binding emission targets under the KP induced a considerable (7–10%) and statistically significant CO_2 reduction effect. According to our results, the KP had no verifiable effect on the CO_2 emissions of ratifying Annex B countries. As a consequence, we are pessimistic that a potential successor of the KP resting on the same principles – even if it would be adopted in the first place – had any discernible effect on the reduction of global GHG emissions. In fact, bilateral or smaller multilateral emission reduction agreements, where commitment is in the best interest of all participating countries, may achieve more (Carbone et al., 2009).

5 Conclusion

The Kyoto Protocol (KP) has been widely criticized by the public press and the scientific community alike. In particular, issues concerning equity, efficiency and cost-effectiveness have been raised. In this paper, we asked in how far the KP lived up to its primary goal, the reduction of domestic GHG emissions in the industrialized world. To answer this question, we analyzed the development of CO_2 emissions for major GHG emitters with binding emission targets under the KP by employing the synthetic control method. We find very little evidence for a significant emission reduction effect for all 15 investigated countries, i.e. countries with binding emissions targets did not emit less CO_2 over the period from 1998–2010 than they would have had were they not subject to GHG emission targets under the KP.

More general, we argue that the empirical challenges faced in the present paper apply to many international environmental policies that can only be evaluated at the country level. Countries tend to be highly heterogeneous with significant differences in their (socio-)economic and political characteristics. As a result, there are numerous issues that need to be addressed, as discussed in Sections 2 and 3. Accounting for all of these challenges simultaneously can be difficult when employing standard panel data analysis. For example, in many cases one has to deal with non-random selection into treatment, a violation of the common trend assumption and the presence of bad controls among important covariates. The synthetic matching method may be better suited to address these challenges than standard panel data approaches.

References

Abadie, A., A. Diamond, and J. Hainmueller (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association 105*, 493–505.

Abadie, A., A. Diamond, and J. Hainmueller (2015). Comparative politics and the synthetic control method. *American Journal of Political Science 59*(2), 495–510.

Abadie, A. and J. Gardeazabal (2003). The economic costs of conflict: A case study of the basque country. *American Economic Review 93*, 113–32.

Aichele, R. and G. Felbermayr (2012). Kyoto and the carbon footprint of nations. *Journal of Environmental Economics and Management 63*, 336–54.

Aichele, R. and G. Felbermayr (2013). The effect of the Kyoto Protocol on carbon emissions. *Journal of Policy Analysis and Management 32*, 731–57.

Angrist, J. and J.-S. Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Barrett, S. (1998). Political economy of the Kyoto Protocol. *Oxford Review of Economic Policy 14*, 20–39.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics 119*, 249–75.

Carbone, J. C., C. Helm, and T. F. Rutherford (2009). The case for international emission trade in the absence of cooperative climate policy. *Journal of Environmental Economics and Management 58*, 266–280.

Cavallo, E., S. Galiani, I. Noy, and J. Pantano (2013, 2015/06/05). Catastrophic natural disasters and economic growth. *Review of Economics and Statistics 95*(5), 1549–1561.

Copeland, B. R. and M. S. Taylor (2005). Free trade and global warming: a trade theory view of the Kyoto Protocol. *Journal of Environmental Economics and Management 49*, 205–34.

Finus, M. (2008). Game theoretic research on the design of international environmental agreements: Insights, critical remarks, and future challenges. *International Review of Environmental and Resource Economics 2*, 29–67.

Grunewald, N. and I. Martinez-Zarzoso (2015, 3). Did the kyoto protocol fail? an evaluation of the effect of the kyoto protocol on co2 emissions. *Environment and Development Economics FirstView*, 1–22.

Imbens, G. W. and J. M. Wooldridge (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature 47*, 5–86.

Prins, G. and S. Rayner (2007). Time to ditsch Kyoto. *Nature 449*, 973–5.

Rubin, D. B. (1976). Inference and missing data. *Biometrika 63*, 581–92.

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Annals of Statistics 6*, 34–58.

Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association 100*, 322–31.

The Economist (1997). Global warming. Rubbing sleep from their eyes. *The Economist (US edition)*. December 11$^{\text{th}}$ 1997.

Tol, R. S. J. (2000). Kyoto, efficiency, and cost-effectiveness: applications of FUND. *Energy Journal 21*, 131–56.

Zhang, J. and C. Wang (2011). Co-benefits and additionality of the clean development mechanism: An empirical analysis. *Journal of Environmental Economics and Management 62*, 140–54.

A Appendix

A.1 Ratification of the KP as the treatment event

Table 4: Weights for using non-Annex B countries in the donor pool

Australia:	Hong Kong (0.299), South Korea (0.211), USA (0.489)
Austria:	Albania (0.085), Brunei Darussalam (0.026), Macao (0.273), Malta (0.062), Singapore (0.366), USA (0.188),
Belgium:	Albania (0.089), Macao (0.024), Malta (0.067), Singapore (0.382), Uruguay (0.068), USA (0.37)
Canada:	Albania (0.05), Malta (0.249), USA (0.647)
Finland:	Albania (0.103), Macao (0.132), Malta (0.213), Singapore (0.305), Uruguay (0.051), USA (0.196
France:	Albania (0.15), Macao (0.06), Singapore (0.597), USA (0.246)
Germany:	Albania (0.132), Gabon (0.015), Singapore (0.29), USA (0.564)
Great Britain:	Albania (0.13), Gabon (0.018), Macao (0.171), Singapore (0.55), USA (0.131),
Italy:	Albania (0.086), Macao (0.215), Malta (0.18), Singapore (0.415), USA (0.104)
Japan:	Malta (0.306), Singapore (0.212), USA (0.483)
Netherlands:	Albania (0.08), Malta (0.139), Singapore (0.353), USA (0.428)
Norway:	Albania (0.01), Brunei Darussalam (0. 067), Macao (0.054), Malta (0.03), Singapore (0.05), USA (0.788)
Portugal:	Costa Rica (0.182), Macao (0.481), Malta (0.014), Cyprus (0.05), Mauritius (0.041), Thailand (0.049) Tunesia (0.024), Uruguay (0.208)
Spain:	Albania (0.053), Cyprus (0.373), Macao (0.149), Malta (0.283), Singapore (0.039), USA (0.103)
Sweden:	Albania (0.147), Singapore (0.295), USA (0.558)

Note: Country Weights for Results in Table 2. The countries in the donor pool, i.e. countries that may receive positive weights, include all non-Annex B countries with sufficient data except for countries classified as low income according to the World Bank. More specifically, the donor pool/control group consists of (ISO 3 country codes): ALB, ARG, BHR, BLZ, BRA, BRB, BRN, BWA, CHL, CHN, CO, CRI, CYP, DOM, ECU, FJI, GAB, IRN, JOR, KOR, MAC, MEX, MLT, MUS, MYS, PAN, PER, SAU, SGP, THA, TTO, TUN, TUR, URY, USA, VEN, and ZAF.

Figure 5: Actual and synthetic CO_2 emissions for non-Annex B countries in donor pool

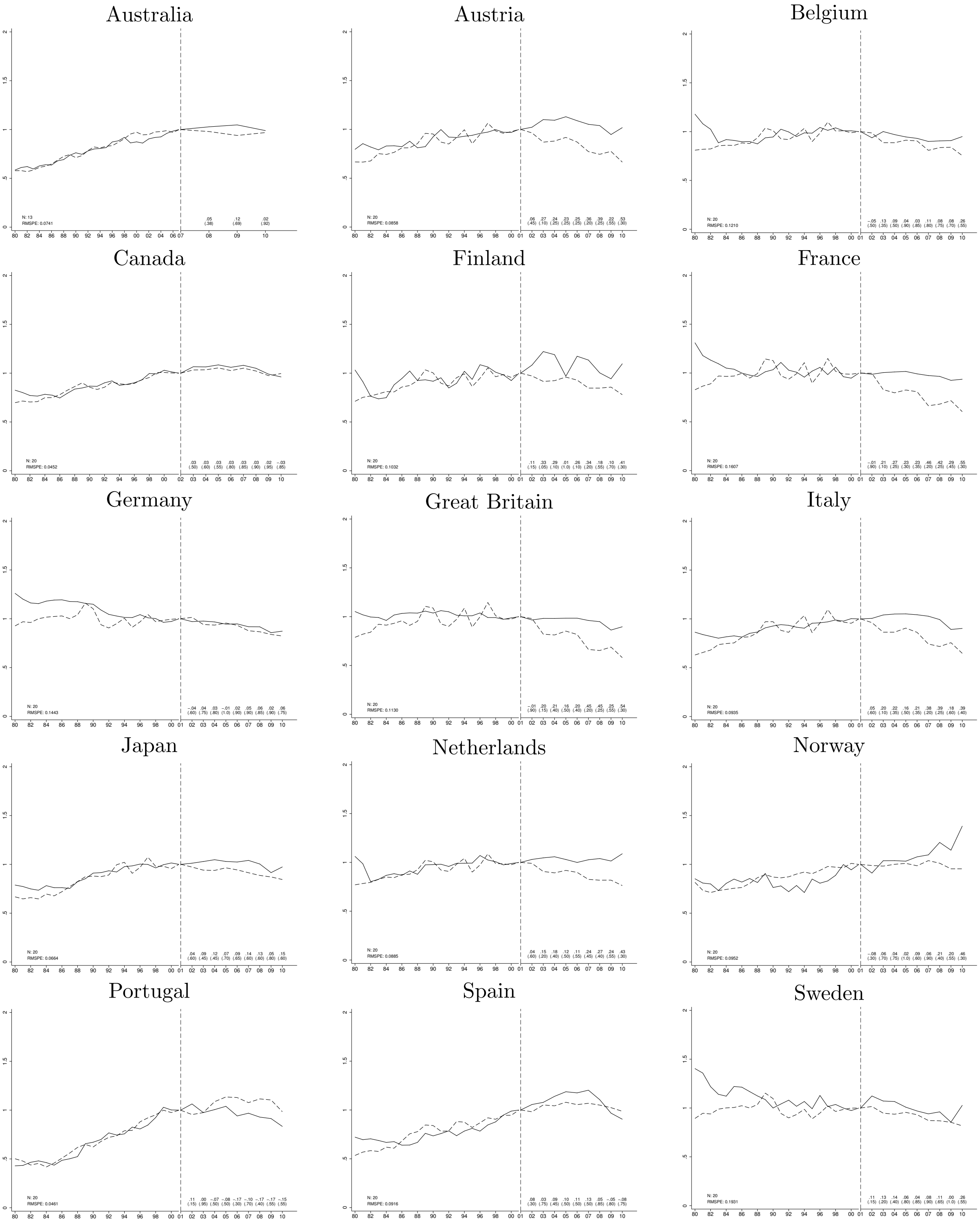


Figure 6: Placebo studies for non-Annex B countries in donor pool

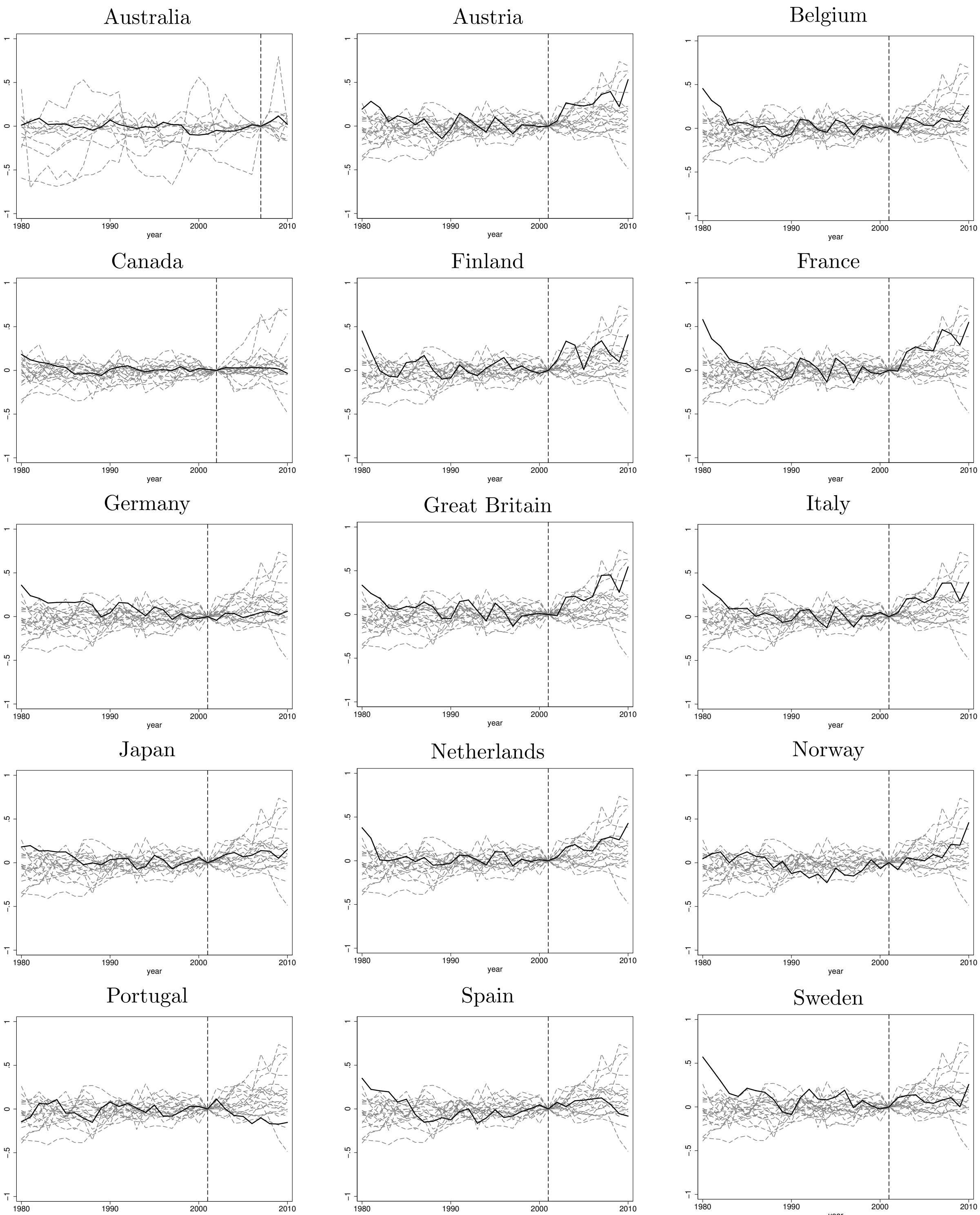


Table 5: Weights for using US states in the Donor Pool

Australia:	Arizona (0.224), Nevada (0.776)
Austria:	Delaware (0.48), Nevada (0.486), Texas (0.034)
Belgium:	Delaware (0.636), Nevada (0.364)
Canada:	Delaware (0.101), District of Columbia (0.171), Nevada (0.727)
Finland:	Delaware (0.541), Nevada (0.418), Texas (0.042)
France:	Arizona (0.007), Delaware (0.868), Nevada (0.125)
Germany:	District of Columbia (0.82), Ohio (0.18)
Great Britain:	Delaware (0.773), Nevada (0.227)
Italy:	Delaware (0.454), Nevada (0.546)
Japan:	Idaho (0.093), Loisiana (0.146), Michigan (0.176), North Dakota (0.033), Oregon (0.004), Rhode Island (0.112), South Dakota (0.323), Utah (0.101)
Netherlands:	Delaware (0.53), Nevada (0.21), Texas (0.26)
Norway:	Delaware (0.291), Nevada (0.709)
Portugal:	Nevada (1.0)
Spain:	Delware (0.071), Nevada (0.929)
Sweden:	Arizona (0.029), Delaware (0.629), District of Columbia (0.342),

Note: US state Weights for Results in Table 3. All US states are included in the donor pool, i.e. may receive positive weights. Given the importance of Delaware and Nevada for many Annex B countries we performed a robustness check excluding both states from the donor pool. Results are almost identical with no evidence for a permament emission reduction effect.

Figure 7: Actual and synthetic CO_2 emissions for US states in donor pool

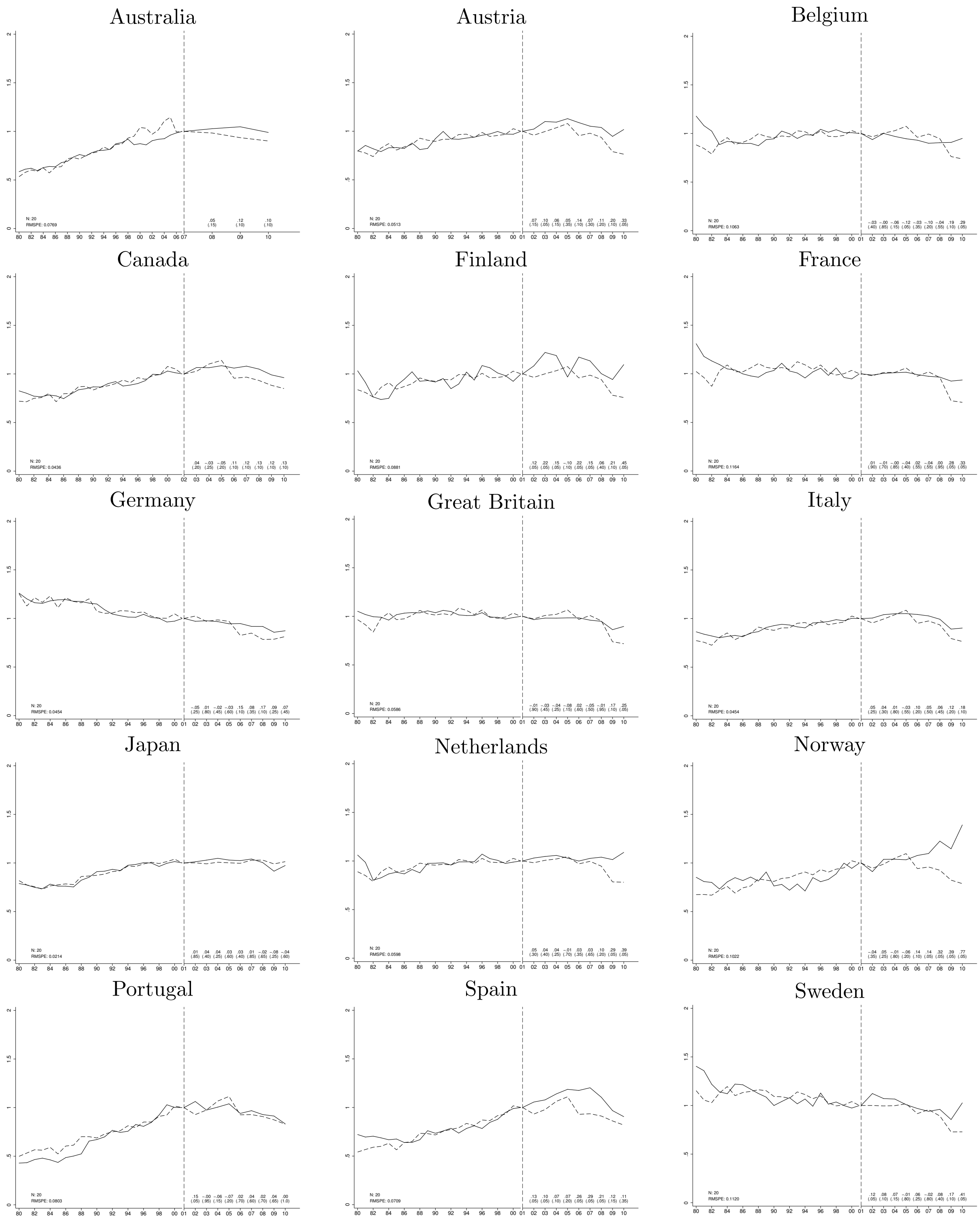
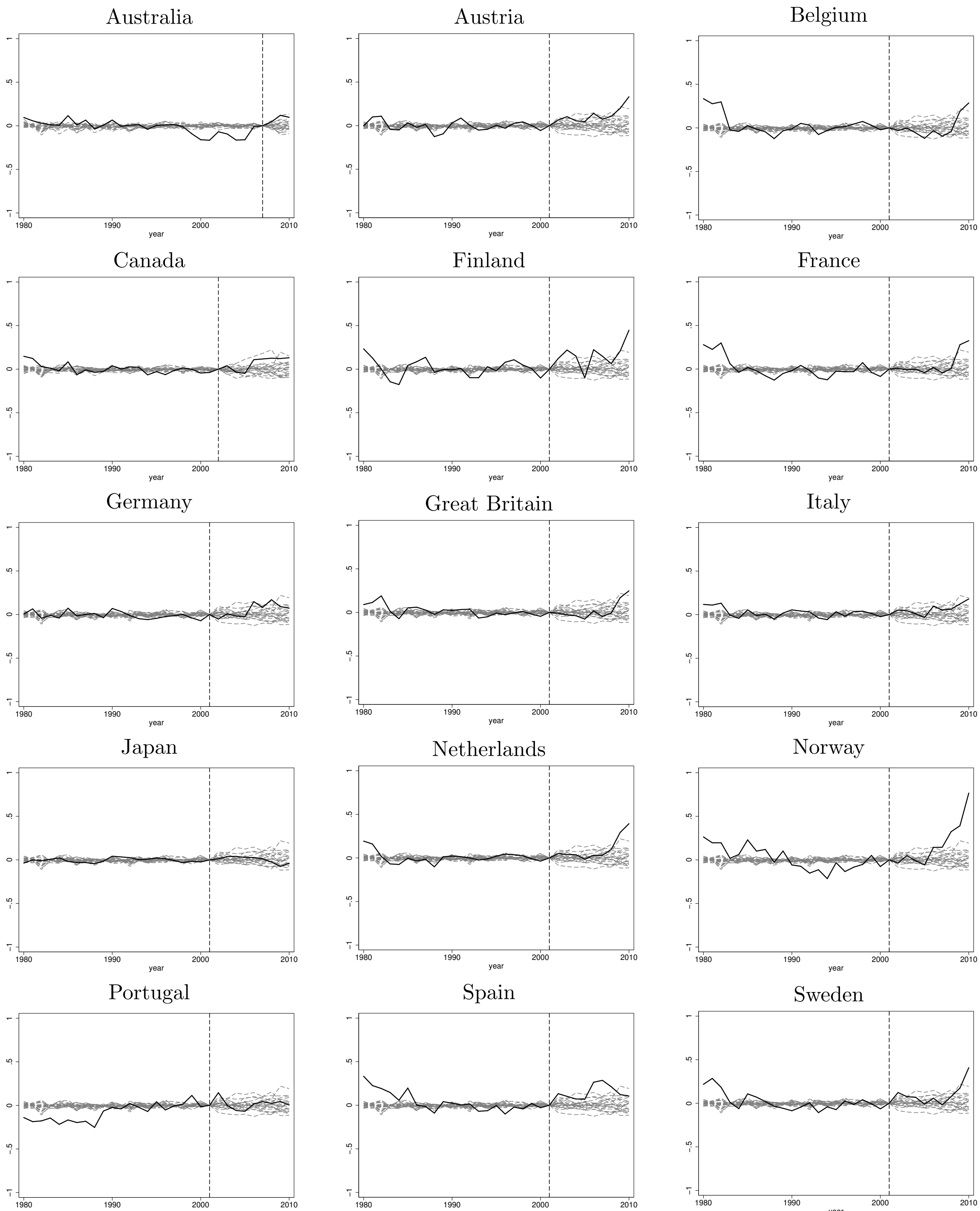


Figure 8: Placebo studies for US states in donor pool



A.2 Adoption of the KP as the treatment event

Table 6: Treatment in 1997: Estimates for Treatment Effects based on non-Annex B countries

Country	Year												
	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Australia	.03	-.05	-.07	-.08	-.04	-.06	-.05	-.02	.00	-.01	.03	.11	.05
	(.75)	(.70)	(.55)	(.65)	(.80)	(.70)	(.80)	(.80)	(1.0)	(1.0)	(.85)	(.70)	(.95)
	RMSPE: 0.0350												
Austria	.06	.01	-.01	.01	.05	.14	.15	.16	.11	.03	-.05	-.10	.04
	(.60)	(.95)	(1.0)	(.90)	(.75)	(.55)	(.45)	(.50)	(.60)	(1.0)	(.85)	(.75)	(1.0)
	RMSPE: 0.0867												
Belgium	.05	-.04	-.01	-.01	-.07	.01	-.04	-.08	-.11	-.10	-.16	-.13	-.03
	(.65)	(.85)	(.95)	(.90)	(.70)	(1.0)	(.80)	(.60)	(.60)	(.85)	(.60)	(.70)	(1.0)
	RMSPE: 0.1064												
Canada	.12	.10	.13	.10	.09	.20	.19	.20	.19	.25	.24	.21	.19
	(.25)	(.45)	(.30)	(.55)	(.70)	(.40)	(.30)	(.35)	(.45)	(.45)	(.50)	(.70)	(.75)
	RMSPE: 0.0485												
Finland	-.01	-.08	-.12	-.07	.03	.16	.12	-.11	.07	-.00	-.17	-.19	.00
	(.80)	(.45)	(.30)	(.70)	(.80)	(.55)	(.50)	(.55)	(.65)	(1.0)	(.55)	(.60)	(1.0)
	RMSPE: 0.0910												
France	.19	.04	.03	.06	.03	.15	.19	.16	.17	.28	.26	.19	.30
	(.10)	(.85)	(.70)	(.85)	(.80)	(.55)	(.35)	(.55)	(.50)	(.55)	(.55)	(.70)	(.65)
	RMSPE: 0.1807												
Germany	.06	-.06	-.04	-.05	-.10	-.09	-.08	-.13	-.09	-.10	-.09	-.12	-.11
	(.60)	(.55)	(.65)	(.80)	(.70)	(.65)	(.60)	(.55)	(.65)	(.85)	(.80)	(.70)	(.75)
	RMSPE: 0.1193												
Great Britain	.00	-.06	-.06	-.06	-.09	-.07	-.06	-.08	-.08	-.14	-.21	-.26	-.17
	(.95)	(.55)	(.60)	(.80)	(.70)	(.70)	(.80)	(.60)	(.65)	(.55)	(.40)	(.40)	(.65)
	RMSPE: 0.0998												
Italy	.08	.06	.06	.03	.04	.10	.17	.13	.14	.10	.03	-.06	.03
	(.30)	(.65)	(.60)	(.85)	(.80)	(.60)	(.40)	(.55)	(.60)	(.85)	(.85)	(.90)	(1.0)
	RMSPE: 0.0859												
Japan	.04	.05	.09	.03	.06	.08	.10	.06	.08	.11	.10	.04	.11
	(.70)	(.75)	(.50)	(.90)	(.75)	(.70)	(.60)	(.70)	(.65)	(.85)	(.80)	(.95)	(.75)
	RMSPE: 0.0590												
Netherlands	-.03	-.09	-.10	-.09	-.08	-.09	-.07	-.11	-.14	-.13	-.14	-.12	-.03
	(.75)	(.45)	(.30)	(.55)	(.70)	(.60)	(.70)	(.55)	(.50)	(.55)	(.60)	(.70)	(1.0)
	RMSPE: 0.0778												
Norway	.08	.20	.07	.15	.04	.20	.19	.18	.26	.17	.32	.32	.59
	(.35)	(.20)	(.55)	(.50)	(.75)	(.35)	(.30)	(.45)	(.30)	(.55)	(.35)	(.30)	(.10)
	RMSPE: 0.0680												
Portugal	.07	.16	.14	.12	.18	.09	.06	.07	-.08	-.12	-.23	-.19	-.23
	(.45)	(.25)	(.25)	(.50)	(.40)	(.65)	(.80)	(.60)	(.65)	(.55)	(.40)	(.50)	(.55)
	RMSPE: 0.0593												
Spain	.17	.20	.31	.22	.31	.36	.46	.47	.49	.56	.46	.29	.24
	(.10)	(.20)	(.10)	(.25)	(.25)	(.20)	(.10)	(.05)	(.10)	(.10)	(.25)	(.50)	(.65)
	RMSPE: 0.1025												
Sweden	.04	-.07	-.11	-.10	-.02	-.08	-.08	-.14	-.15	-.18	-.16	-.22	-.06
	(.70)	(.45)	(.30)	(.50)	(.85)	(.65)	(.60)	(.50)	(.45)	(.55)	(.60)	(.55)	(.95)
	RMSPE: 0.1608												

Note: The Table contains treatment effect estimates for each Annex B country under investigation using non-Annex B countries to construct the synthetic counterfactuals and considering the year 1997 (the year of adoption of the KP) as the time of treatment. For each country, we display the yearly treatment effects in percent and the respective probabilities of finding such an effect in parenthesis. Finally, we report the (pre-treatment) RMSPE for each country.

Figure 9: Treatment in 1997: Actual and synthetic CO_2 emissions using non-Annex B countries in the donor pool

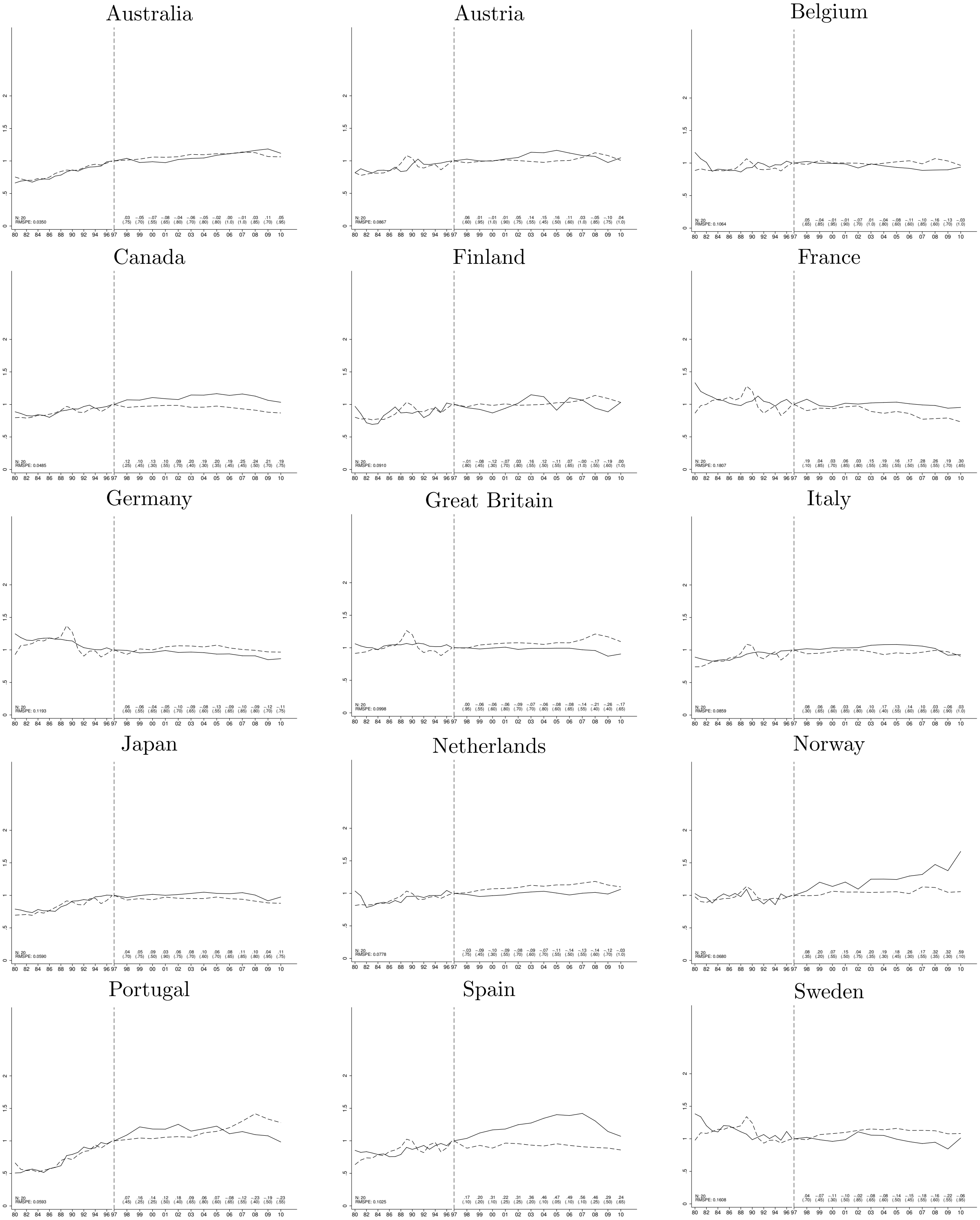
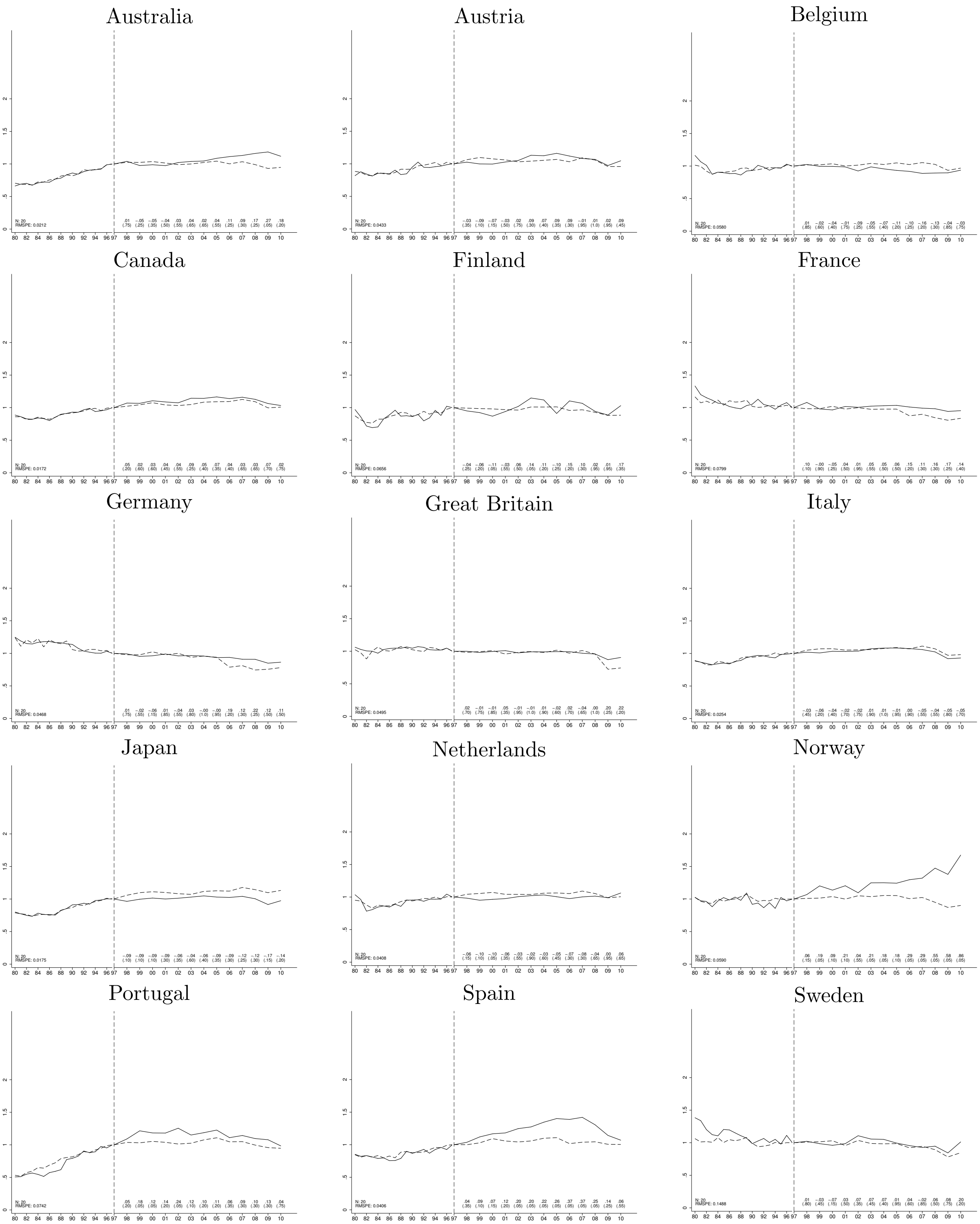


Table 7: Treatment in 1997: Estimates for Treatment Effects based on US states

Country	Year 1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Australia	.01	-.05	-.05	-.04	.03	.04	.02	.04	.11	.09	.17	.27	.18
	(.75)	(.25)	(.35)	(.50)	(.55)	(.65)	(.65)	(.55)	(.25)	(.30)	(.25)	(.05)	(.20)
	RMSPE: 0.0212												
Austria	-.03	-.09	-.07	-.03	.02	.09	.07	.09	.09	-.01	.01	.02	.09
	(.35)	(.10)	(.15)	(.50)	(.75)	(.30)	(.40)	(.35)	(.30)	(.95)	(1.0)	(.95)	(.45)
	RMSPE: 0.0433												
Belgium	.01	-.02	-.04	-.01	-.09	-.05	-.07	-.11	-.10	-.16	-.13	-.04	-.03
	(.85)	(.60)	(.40)	(.75)	(.25)	(.55)	(.40)	(.20)	(.25)	(.20)	(.30)	(.85)	(.75)
	RMSPE: 0.0580												
Canada	.05	.02	.03	.04	.04	.09	.05	.07	.04	.03	.03	.07	.02
	(.20)	(.60)	(.60)	(.45)	(.55)	(.25)	(.40)	(.35)	(.40)	(.65)	(.65)	(.70)	(.75)
	RMSPE: 0.0172												
Finland	-.04	-.06	-.11	-.03	.06	.14	.11	-.10	.15	.10	.02	.01	.17
	(.25)	(.20)	(.05)	(.55)	(.50)	(.05)	(.20)	(.25)	(.20)	(.30)	(.95)	(.95)	(.35)
	RMSPE: 0.0656												
France	.10	-.00	-.05	.04	.01	.05	.05	.06	.15	.11	.16	.17	.14
	(.10)	(.90)	(.25)	(.50)	(.95)	(.55)	(.50)	(.50)	(.20)	(.30)	(.30)	(.25)	(.40)
	RMSPE: 0.0799												
Germany	.01	-.02	-.06	.01	-.04	.03	-.00	-.00	.19	.12	.22	.12	.11
	(.75)	(.55)	(.15)	(.85)	(.55)	(.80)	(1.0)	(.95)	(.20)	(.30)	(.25)	(.50)	(.50)
	RMSPE: 0.0468												
Great Britain	.02	-.01	-.01	.05	-.01	-.01	.01	-.02	.02	-.04	.00	.20	.22
	(.70)	(.75)	(.85)	(.35)	(.95)	(1.0)	(.90)	(.60)	(.70)	(.65)	(1.0)	(.25)	(.20)
	RMSPE: 0.0495												
Italy	-.03	-.06	-.04	-.02	-.02	.01	.01	-.01	.00	-.05	-.04	-.05	-.05
	(.45)	(.20)	(.40)	(.70)	(.75)	(.90)	(1.0)	(.95)	(.90)	(.55)	(.55)	(.80)	(.70)
	RMSPE: 0.0254												
Japan	-.09	-.09	-.09	-.09	-.06	-.04	-.06	-.09	-.09	-.12	-.12	-.17	-.14
	(.10)	(.10)	(.10)	(.30)	(.35)	(.60)	(.40)	(.35)	(.30)	(.25)	(.30)	(.15)	(.20)
	RMSPE: 0.0175												
Netherlands	-.06	-.10	-.10	-.06	-.03	-.02	-.03	-.05	-.07	-.08	-.04	.00	.06
	(.15)	(.10)	(.05)	(.35)	(.55)	(.90)	(.60)	(.45)	(.30)	(.30)	(.65)	(.95)	(.65)
	RMSPE: 0.0408												
Norway	.06	.19	.09	.21	.04	.21	.18	.18	.29	.29	.55	.58	.86
	(.15)	(.05)	(.10)	(.10)	(.55)	(.05)	(.05)	(.10)	(.05)	(.05)	(.05)	(.05)	(.05)
	RMSPE: 0.0590												
Portugal	.05	.18	.12	.14	.24	.12	.10	.11	.06	.09	.10	.13	.04
	(.20)	(.05)	(.05)	(.20)	(.05)	(.10)	(.20)	(.20)	(.35)	(.30)	(.30)	(.30)	(.75)
	RMSPE: 0.0742												
Spain	.04	.09	.07	.12	.20	.20	.22	.26	.37	.37	.25	.14	.06
	(.35)	(.10)	(.15)	(.20)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.25)	(.55)
	RMSPE: 0.0406												
Sweden	.01	-.03	-.07	.03	.07	.07	.07	.01	.04	-.02	.06	.08	.20
	(.80)	(.45)	(.15)	(.50)	(.35)	(.45)	(.40)	(.95)	(.60)	(.85)	(.50)	(.75)	(.20)
	RMSPE: 0.1488												

Note: The Table contains treatment effect estimates for each Annex B country under investigation using US states to construct the synthetic counterfactuals and considering the year 1997 (the year of adoption of the KP) as the time of treatment. For each country, we display the yearly treatment effects in percent and the respective probabilities of finding such an effect in parenthesis. Finally, we report the (pre-treatment) RMSPE for each country.

Figure 10: Treatment in 1997: Actual and synthetic CO_2 emissions using US states in the donor pool



A.3 KP's entering into force as the treatment event

Table 8: Treatment in 2004: Estimates for Treatment Effects based on non-Annex B Countries

Country	non-Annex B countries						US states					
	Year						Year					
	2005	2006	2007	2008	2009	2010	2005	2006	2007	2008	2009	2010
Australia	-.01	.01	.01	.04	.12	.07	.04	.08	.08	.14	.23	.14
	(.95)	(1.0)	(.95)	(.75)	(.65)	(.85)	(.40)	(.10)	(.20)	(.05)	(.05)	(.10)
	RMSPE: 0.0372						RMSPE: 0.0237					
Austria	.10	.09	-.01	-.03	-.08	.01	.08	.06	.01	.02	.05	.08
	(.45)	(.50)	(.95)	(.85)	(.75)	(1.0)	(.15)	(.20)	(.85)	(.70)	(.65)	(.25)
	RMSPE: 0.0948						RMSPE: 0.0436					
Belgium	-.08	-.08	-.09	-.08	-.06	.01	-.11	-.11	-.16	-.10	-.02	-.02
	(.55)	(.55)	(.70)	(.75)	(.80)	(1.0)	(.10)	(.05)	(.05)	(.15)	(.80)	(.80)
	RMSPE: 0.1215						RMSPE: 0.0551					
Canada	.13	.10	.13	.10	.08	.08	-.04	-.05	-.05	-.03	-.01	-.05
	(.35)	(.45)	(.55)	(.65)	(.75)	(.85)	(.20)	(.20)	(.30)	(.65)	(.95)	(.60)
	RMSPE: 0.0617						RMSPE: 0.0228					
Finland	-.12	.07	-.01	-.14	-.16	.01	-.19	.02	-.01	-.10	-.07	.07
	(.35)	(.55)	(.95)	(.50)	(.50)	(1.0)	(.05)	(.60)	(.85)	(.15)	(.40)	(.50)
	RMSPE: 0.0962						RMSPE: 0.0664					
France	.09	.08	.15	.10	.05	.16	.06	.15	.10	.16	.18	.15
	(.55)	(.55)	(.55)	(.75)	(.80)	(.60)	(.20)	(.05)	(.20)	(.10)	(.10)	(.15)
	RMSPE: 0.1589						RMSPE: 0.0728					
Germany	-.09	-.05	-.04	-.04	-.09	-.05	-.01	.19	.12	.22	.12	.10
	(.50)	(.65)	(.80)	(.85)	(.75)	(.90)	(.95)	(.05)	(.10)	(.05)	(.20)	(.25)
	RMSPE: 0.1089						RMSPE: 0.0426					
Great Britain	.00	.04	.01	-.01	-.09	-.01	-.02	.07	-.00	.04	.07	.10
	(.95)	(.70)	(.95)	(.90)	(.75)	(1.0)	(.75)	(.20)	(1.0)	(.65)	(.45)	(.25)
	RMSPE: 0.1013						RMSPE: 0.0293					
Italy	.12	.13	.14	.09	-.02	.06	.01	.02	-.02	-.02	-.03	-.05
	(.40)	(.40)	(.55)	(.75)	(.95)	(.90)	(.85)	(.60)	(.85)	(.75)	(.75)	(.60)
	RMSPE: 0.0888						RMSPE: 0.0206					
Japan	-.01	.01	.01	.00	-.05	.00	-.03	-.05	-.06	-.06	-.10	-.08
	(.90)	(1.0)	(.95)	(.95)	(.80)	(1.0)	(.45)	(.25)	(.30)	(.25)	(.15)	(.25)
	RMSPE: 0.0606						RMSPE: 0.0219					
Netherlands	-.05	-.07	-.04	-.03	-.02	.08	-.03	-.06	-.06	.00	.07	.09
	(.65)	(.55)	(.80)	(.85)	(.95)	(.85)	(.45)	(.20)	(.30)	(.95)	(.40)	(.25)
	RMSPE: 0.0795						RMSPE: 0.0353					
Norway	.17	.26	.22	.38	.36	.64	.05	.12	.12	.28	.27	.52
	(.25)	(.20)	(.40)	(.15)	(.25)	(.10)	(.20)	(.05)	(.05)	(.05)	(.05)	(.05)
	RMSPE: 0.0953						RMSPE: 0.0706					
Portugal	.08	-.06	-.09	-.19	-.16	-.19	-.00	.01	.03	.01	.04	-.01
	(.55)	(.55)	(.60)	(.40)	(.50)	(.40)	(1.0)	(.75)	(.65)	(.90)	(.65)	(.85)
	RMSPE: 0.0857						RMSPE: 0.0901					
Spain	.35	.33	.36	.24	.11	.10	.10	.07	.08	-.00	-.05	-.13
	(.05)	(.10)	(.15)	(.40)	(.70)	(.75)	(.05)	(.05)	(.05)	(.90)	(.40)	(.05)
	RMSPE: 0.1218						RMSPE: 0.0395					
Sweden	-.13	-.15	-.17	-.16	-.22	-.05	.05	.17	.11	.22	.09	.26
	(.30)	(.30)	(.45)	(.45)	(.35)	(.85)	(.30)	(.05)	(.20)	(.05)	(.35)	(.05)
	RMSPE: 0.1596						RMSPE: 0.0823					

Note: The Table contains treatment effect estimates for each Annex B country under investigation using non-Annex B countries and US states to construct the synthetic counterfactuals and considering the year 2004 (when the KP entered into force) as the time of treatment. For each country, we display the yearly treatment effects in percent and the respective probabilities of finding such an effect in parenthesis. Finally, we report the (pre-treatment) RMSPE for each country.

Figure 11: Treatment in 2004: Actual and synthetic CO_2 emissions for non-Annex B countries in donor pool

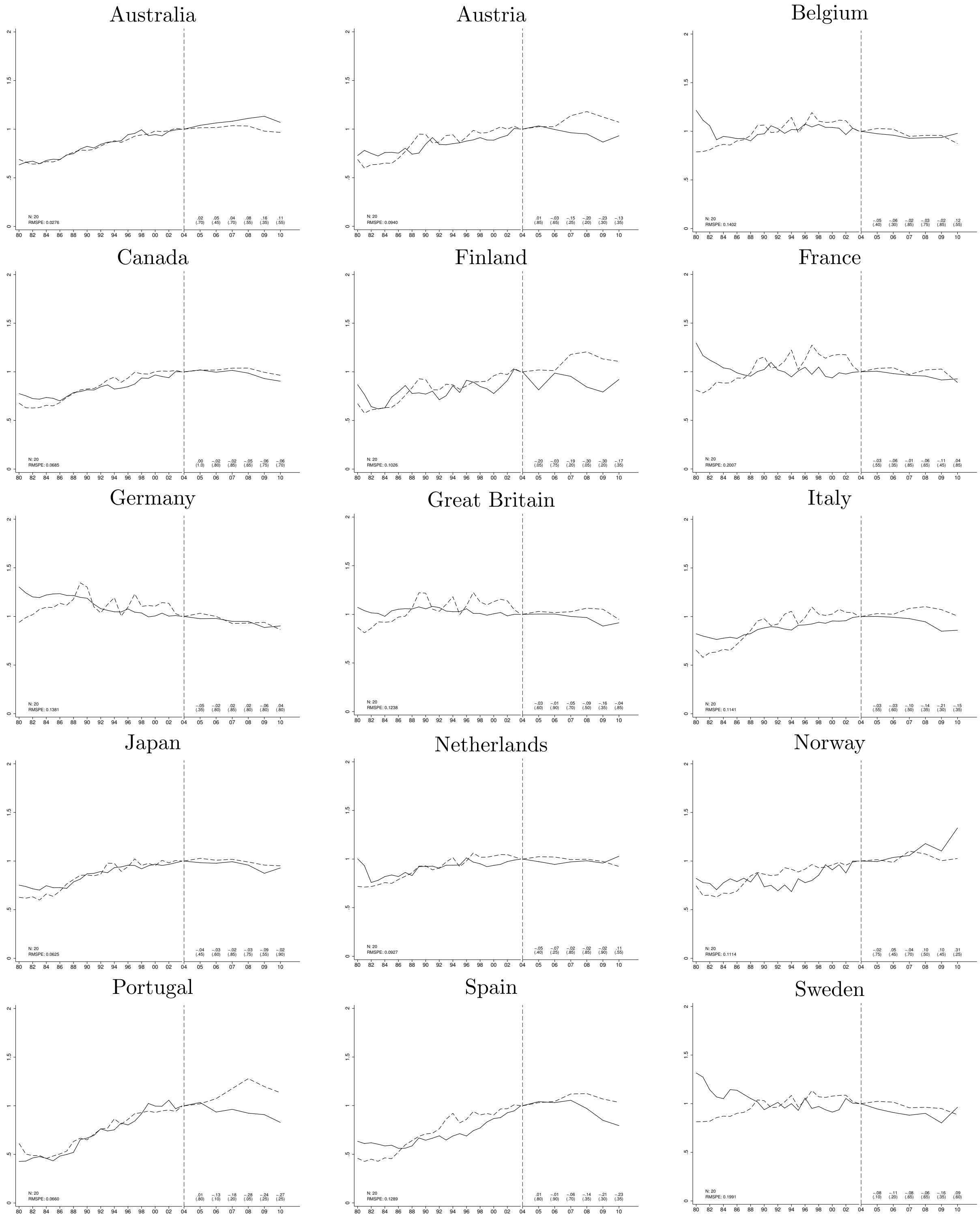
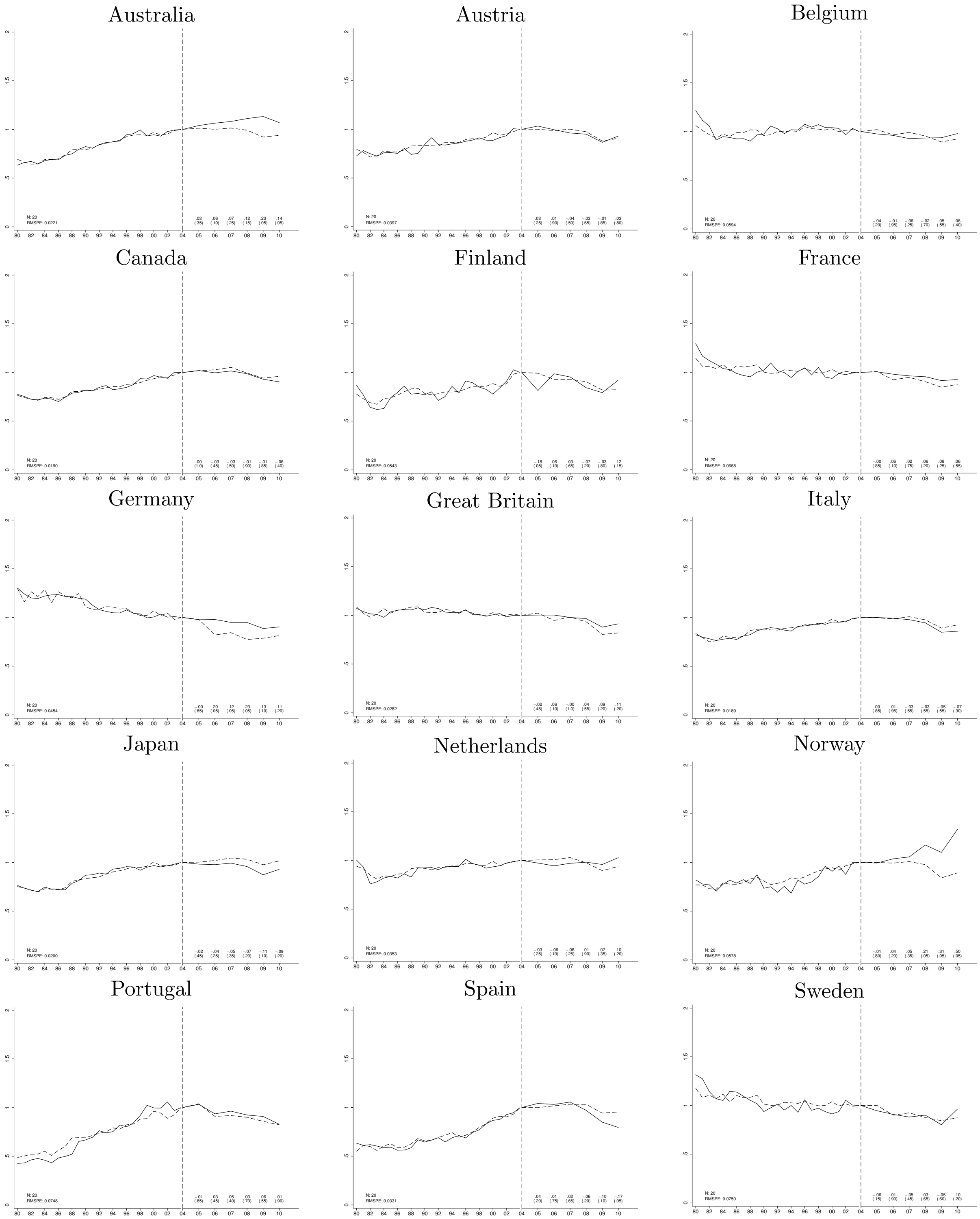


Figure 12: Treatment in 2004: Actual and synthetic CO_2 emissions for US states in donor pool



A.4 CO_2 intensity as dependent variable

Table 9: Estimates for Treatment Effects (CO_2 per GDP) based on non-Annex B Countries

Country	Year 2002	2003	2004	2005	2006	2007	2008	2009	2010
Australia							.08	.08	.03
							(.45)	(.60)	(.85)
	RMSPE: 0.2353								
Austria	.07	.19	.20	.29	.31	.19	.13	.15	.22
	(.45)	(.25)	(.20)	(.20)	(.35)	(.35)	(.70)	(.50)	(.55)
	RMSPE: 0.1479								
Belgium	-.05	.05	.04	.04	.12	.15	.18	.17	.27
	(.55)	(.80)	(.85)	(.80)	(.60)	(.55)	(.60)	(.50)	(.50)
	RMSPE: 0.1085								
Canada		.03	.04	.06	.06	.05	.07	.04	.01
		(.75)	(.65)	(.70)	(.80)	(.90)	(.75)	(.65)	(.95)
	RMSPE: 0.0760								
Finland	.15	.29	.25	.05	.33	.19	.04	.11	.29
	(.20)	(.15)	(.15)	(.75)	(.30)	(.35)	(.80)	(.55)	(.40)
	RMSPE: 0.1536								
France	.01	.19	.30	.39	.46	.63	.66	.59	.81
	(.80)	(.30)	(.15)	(.20)	(.25)	(.20)	(.25)	(.25)	(.15)
	RMSPE: 0.1616								
Germany	-.05	-.06	-.11	-.13	-.11	-.15	-.14	-.11	-.11
	(.55)	(.70)	(.45)	(.45)	(.55)	(.40)	(.60)	(.55)	(.65)
	RMSPE: 0.1769								
Great Britain	-.00	.10	.10	.19	.25	.35	.36	.30	.48
	(1.0)	(.65)	(.50)	(.35)	(.35)	(.30)	(.35)	(.45)	(.25)
	RMSPE: 0.1284								
Italy	.14	.26	.35	.45	.50	.44	.33	.34	.44
	(.20)	(.25)	(.15)	(.20)	(.15)	(.25)	(.35)	(.35)	(.30)
	RMSPE: 0.1522								
Japan	.07	.09	.12	.15	.18	.21	.21	.21	.26
	(.45)	(.65)	(.45)	(.45)	(.35)	(.35)	(.55)	(.50)	(.50)
	RMSPE: 0.1735								
Netherlands	.05	.13	.18	.12	.14	.22	.27	.29	.42
	(.55)	(.50)	(.25)	(.50)	(.55)	(.35)	(.45)	(.40)	(.25)
	RMSPE: 0.1435								
Norway	-.04	.10	.07	.07	.18	.11	.29	.15	.44
	(.70)	(.60)	(.65)	(.70)	(.35)	(.60)	(.30)	(.50)	(.15)
	RMSPE: 0.2199								
Portugal	.08	.04	.08	.09	-.04	.02	-.14	-.09	-.11
	(.30)	(.80)	(.55)	(.55)	(.95)	(.95)	(.60)	(.60)	(.65)
	RMSPE: 0.0935								
Spain	.08	.18	.30	.43	.40	.51	.40	.23	.28
	(.35)	(.35)	(.15)	(.20)	(.30)	(.20)	(.35)	(.50)	(.55)
	RMSPE: 0.1359								
Sweden	.09	.05	.01	.01	-.00	-.03	.01	-.02	.15
	(.20)	(.80)	(.90)	(.90)	(.95)	(.95)	(.85)	(.90)	(.60)
	RMSPE: 0.2353								

Note: The Table contains treatment effect estimates for CO_2 emissions per GDP for each Annex B country under investigation using non-Annex B countries to construct the synthetic counterfactuals and considering the year of ratification as the time of treatment. For each country, we display the yearly treatment effects in percent and the respective probabilities of finding such an effect in parenthesis. Finally, we report the (pre-treatment) RMSPE for each country.

A.5 CO_2 per capita as dependent variable

Table 10: Estimates for Treatment Effects (CO_2 per capita) based on non-Annex B Countries

Country	Year 2002	2003	2004	2005	2006	2007	2008	2009	2010
Australia							.04	.09	-.00
							(.55)	(.45)	(1.0)
	RMSPE: 0.0446								
Austria	.09	.19	.15	.14	.19	.22	.30	.19	.41
	(.25)	(.05)	(.30)	(.25)	(.15)	(.40)	(.30)	(.45)	(.30)
	RMSPE: 0.0688								
Belgium	-.03	.06	.01	-.04	-.01	.01	.06	.07	.18
	(.65)	(.55)	(.90)	(.80)	(.95)	(1.0)	(.85)	(.70)	(.55)
	RMSPE: 0.1090								
Canada		.04	.03	.03	.04	.03	.03	.00	-.05
		(.35)	(.55)	(.50)	(.70)	(.80)	(.85)	(1.0)	(.75)
	RMSPE: 0.0560								
Finland	.14	.27	.21	-.05	.23	.24	.17	.11	.35
	(.10)	(.05)	(.15)	(.75)	(.15)	(.35)	(.50)	(.65)	(.30)
	RMSPE: 0.0967								
France	.01	.13	.16	.12	.15	.29	.31	.24	.41
	(.95)	(.25)	(.30)	(.40)	(.35)	(.35)	(.35)	(.45)	(.30)
	RMSPE: 0.1451								
Germany	-.02	.03	.04	-.00	.04	.06	.09	.07	.11
	(.70)	(.70)	(.75)	(1.0)	(.70)	(.80)	(.70)	(.75)	(.55)
	RMSPE: 0.0995								
Great Britain	.01	.12	.10	.06	.14	.29	.36	.21	.43
	(.80)	(.25)	(.35)	(.65)	(.45)	(.35)	(.30)	(.45)	(.30)
	RMSPE: 0.0768								
Italy	.07	.12	.10	.06	.13	.21	.25	.11	.25
	(.35)	(.25)	(.35)	(.60)	(.50)	(.45)	(.45)	(.65)	(.35)
	RMSPE: 0.0720								
Japan	.05	.04	.07	.03	.07	.08	.09	.04	.11
	(.35)	(.70)	(.50)	(.90)	(.60)	(.75)	(.70)	(.75)	(.55)
	RMSPE: 0.0681								
Netherlands	.04	.10	.14	.09	.08	.17	.20	.20	.36
	(.55)	(.35)	(.30)	(.45)	(.60)	(.50)	(.50)	(.45)	(.30)
	RMSPE: 0.0758								
Norway	-.07	.05	.03	.02	.10	.05	.21	.20	.44
	(.35)	(.60)	(.75)	(.95)	(.55)	(.85)	(.35)	(.40)	(.25)
	RMSPE: 0.1180								
Portugal	.09	-.03	-.07	-.07	-.12	-.08	-.06	-.06	-.10
	(.15)	(.70)	(.50)	(.45)	(.35)	(.70)	(.75)	(.70)	(.55)
	RMSPE: 0.0702								
Spain	.07	.01	.06	.07	.07	.05	-.02	-.10	-.13
	(.30)	(.85)	(.55)	(.45)	(.60)	(.85)	(.90)	(.65)	(.55)
	RMSPE: 0.1050								
Sweden	.12	.15	.16	.08	.07	.10	.14	.04	.29
	(.10)	(.20)	(.25)	(.45)	(.60)	(.70)	(.55)	(.75)	(.30)
	RMSPE: 0.1468								

Note: The Table contains treatment effect estimates for CO_2 emissions per capita for each Annex B country under investigation using non-Annex B countries to construct the synthetic counterfactuals and considering the year of ratification as the time of treatment. For each country, we display the yearly treatment effects in percent and the respective probabilities of finding such an effect in parenthesis. Finally, we report the (pre-treatment) RMSPE for each country.

Table 11: Estimates for Treatment Effects (CO_2 per capita) based on US States

Country	Year 2002	2003	2004	2005	2006	2007	2008	2009	2010
Australia							.05	.12	.04
							(.20)	(.05)	(.35)
	RMSPE: 0.0614								
Austria	.08	.22	.13	.17	.18	.13	.19	.13	.17
	(.10)	(.05)	(.10)	(.05)	(.05)	(.15)	(.10)	(.15)	(.10)
	RMSPE: 0.0620								
Belgium	-.01	.08	-.01	-.02	.01	-.03	.02	.08	.10
	(.70)	(.10)	(.70)	(.45)	(.80)	(.65)	(.65)	(.50)	(.30)
	RMSPE: 0.0561								
Canada		.06	.00	.02	.03	.05	.05	.04	-.00
		(.05)	(.90)	(.50)	(.50)	(.30)	(.30)	(.70)	(.90)
	RMSPE: 0.0469								
Finland	.13	.29	.19	-.02	.24	.21	.12	.11	.29
	(.05)	(.05)	(.05)	(.45)	(.05)	(.05)	(.25)	(.25)	(.05)
	RMSPE: 0.0820								
France	.08	.08	.03	.02	.17	.17	.20	.21	.27
	(.15)	(.10)	(.40)	(.50)	(.05)	(.10)	(.10)	(.10)	(.05)
	RMSPE: 0.1024								
Germany	-.01	.01	-.01	-.03	-.00	-.05	.02	.10	.16
	(.55)	(.75)	(.75)	(.40)	(.90)	(.45)	(.80)	(.40)	(.15)
	RMSPE: 0.0512								
Great Britain	.04	.07	.01	.01	.14	.12	.16	.10	.16
	(.25)	(.10)	(.60)	(.75)	(.05)	(.25)	(.25)	(.45)	(.15)
	RMSPE: 0.0503								
Italy	.06	.14	.08	.08	.10	.07	.10	.03	.01
	(.20)	(.05)	(.10)	(.15)	(.20)	(.30)	(.30)	(.75)	(.90)
	RMSPE: 0.0540								
Japan	.02	.04	.04	.04	-.01	-.01	-.02	-.05	-.03
	(.45)	(.40)	(.20)	(.30)	(.90)	(.90)	(.65)	(.50)	(.70)
	RMSPE: 0.0187								
Netherlands	.10	.14	.09	.06	.14	.17	.25	.28	.37
	(.10)	(.05)	(.10)	(.15)	(.05)	(.10)	(.05)	(.05)	(.05)
	RMSPE: 0.0650								
Norway	-.04	.14	.07	.07	.13	.13	.34	.30	.50
	(.25)	(.05)	(.10)	(.15)	(.05)	(.10)	(.05)	(.05)	(.05)
	RMSPE: 0.1020								
Portugal	.09	-.01	.04	.03	-.05	-.06	-.09	-.06	-.14
	(.10)	(.85)	(.20)	(.40)	(.30)	(.40)	(.30)	(.50)	(.15)
	RMSPE: 0.1355								
Spain	.10	.17	.15	.19	.19	.18	.16	.05	-.06
	(.10)	(.05)	(.05)	(.05)	(.05)	(.05)	(.15)	(.65)	(.55)
	RMSPE: 0.1510								
Sweden	.22	.15	.08	.03	.14	.12	.17	.11	.36
	(.05)	(.05)	(.10)	(.45)	(.05)	(.25)	(.15)	(.30)	(.05)
	RMSPE: 0.0896								

Note: The Table contains treatment effect estimates for CO_2 emissions per capita for each Annex B country under investigation using US states to construct the synthetic counterfactuals and considering the year of ratification as the time of treatment. For each country, we display the yearly treatment effects in percent and the respective probabilities of finding such an effect in parenthesis. Finally, we report the (pre-treatment) RMSPE for each country.