Grantham Research Institute on Climate Change and the Environment

Have international pollution protocols made a difference?

Elisabeth Isaksen

December 2018

Grantham Research Institute on Climate Change and the Environment
Working Paper No. 310

ISSN 2515-5717 (Online)

The Grantham Foundation for the Protection of the Environment

The London School of Economics and Political Science

The Grantham Research Institute on Climate Change and the Environment was established by the London School of Economics and Political Science in 2008 to bring together international expertise on economics, finance, geography, the environment, international development and political economy to create a world-leading centre for policy-relevant research and training. The Institute is funded by the Grantham Foundation for the Protection of the Environment and the Global Green Growth Institute. It has six research themes:

1. Sustainable development
2. Finance, investment and insurance
3. Changing behaviours
4. Growth and innovation
5. Policy design and evaluation
6. Governance and legislation

More information: www.lse.ac.uk/GranthamInstitute

Suggested citation:
Isaksen E (2018) *Have international pollution protocols made a difference?* Grantham Research Institute on Climate Change and the Environment Working Paper 310. London: London School of Economics and Political Science

Have international pollution protocols made a difference?*

Elisabeth Thuestad Isaksen†‡

December 6, 2018

Abstract

Evaluating the effectiveness of international agreements is inherently difficult due to the problems of self-selection, spillovers, and aggregate-level data. In this paper, I provide new and arguably more credible estimates on the effects of the Long Range Transboundary Air Pollution (LRTAP) protocols on three different pollutants: SO_2, NO_x, and VOCs. I address the problem of non-parallel emission trends by constructing "synthetic" controls that mimic the pre-treatment development of each affected country. Using a new dataset covering more regions and a longer period than previously applied, I find that all three protocols induced emissions reductions well beyond a (synthetic) counterfactual development.

Keywords: international environmental agreements, pollution, emissions, synthetic control method

JEL codes: Q53, Q58, F53

*I want to thank Torben Mideksa, Andreas Lange, Mari Rege, Kjell Arne Brekke, Jo Thori Lind, Daire McCoy, and audiences at Columbia University, University of Oslo, and WCERE 2018 for helpful comments and suggestions. This research was supported with funding from the Grantham Foundation (through the Grantham Research Institute on Climate Change and the Environment), and the Research Council of Norway.

†Grantham Research Institute on Climate Change and the Environment, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, United Kingdom. Email: elisabeth.isaksen@gmail.com.

‡The Ragnar Frisch Centre for Economic Research, Norway.

1 Introduction

A wide range of environmental problems are characterized by cross-border externalities. As unilateral action cannot effectively solve these problems, some form of international cooperation is needed. Air pollution has been a major focus of international environmental agreements, and since the 1970s over 60 multilateral treaties, protocols, and amendments have been put in place to address this issue (Mitchell, 2015). The potential of international agreements to deliver emission reductions is intensively discussed in the economic literature, and the majority of theoretical studies postulate that free-riding incentives will undermine the effectiveness of such voluntary efforts.[1] Empirically validating these predictions, however, is methodologically challenging, and to date there are few studies applying a credible framework for causal inference. The fundamental problem is establishing a credible counterfactual for countries voluntary entering into agreements. How would emissions have evolved in absence of participation? Answering this question is complicated by problems such as self-selection bias, spillovers, and anticipation effects.

In this paper, I examine the 1979 Convention on Long-range Transboundary Air Pollution (LRTAP in the following) and three subsequent protocols with the aim of identifying *causal* effects of the protocols on emissions. The LRTAP framework was the first attempt to deal with problems of air pollution on a broad regional basis, covering countries in Europe and North-America. It was initially conceived as a flexible framework for cooperation, but has later been extended by several protocols containing legally binding targets for emissions reductions. Here, I focus on the first three protocols: the 1985 Helsinki protocol (on SO_2), the 1988 Sofia protocol (on NO_x), and the 1991 Geneva protocol (on VOCs).[2]

While there are no empirical examinations of the Geneva protocol (as far as I am aware), recent studies find no effect of the Helsinki protocol on SO_2 emissions (e.g., Ringquist and Kostadinova, 2005; Naughton, 2010; Aakvik and Tjøtta, 2011), and significant, but small reduc-

[1] The theoretical literature on international environmental agreements is vast. See e.g., Barrett (1994) and Hoel (1992) for seminal papers, and Benchekroun and Long (2012) or Jørgensen et al. (2010) for literature reviews.

[2] SO_2 is short for sulfur dioxides, NO_x is short for nitrogen oxides and VOC is short for volatile organic compounds. All pollutants are associated with adverse health effects, and can travel long distances before depositing and causing damage to ecosystems such as forests and lakes. See Appendix A.2 for more details.

tions in NO_x emissions induced by the Sofia protocol (Bratberg et al., 2005; Naughton, 2010).[3] Causal interpretation of these findings rely on several identifying assumptions that I argue are not adequately addressed in the studies – in large part due to data limitations and methodological choices. Specifically, the development in pre-treatment emissions is likely to correlate with participation, treatment effects may spill to nearby control countries, and anticipation effects might materialize before the formal ratification of the protocol.

The main contribution of this paper is to provide new and arguably more credible causal evidence on the effects of the Helsinki, Sofia, and Geneva protocols on emissions, by combining a new global dataset on emissions with a relatively recent methodology to construct counterfactual developments: the synthetic control method (Abadie and Gardeazabal, 2003; Abadie et al., 2010). The method was initially developed as a data-driven procedure to construct a suitable counterfactual in cases with few treated units, with the underlying idea that a weighted combination of control countries likely serves as a better comparison than any single country alone. The "synthetic" control country is constructed by assigning weights to plausibly unaffected countries, where the weights are chosen on the basis of how well the synthetic control approximates the development in important pre-treatment variables, such as past emissions.

The method requires data on a sufficiently long pre-intervention period and a large *donor pool* of potential control countries. Previous studies have almost exclusively relied on the officially reported LRTAP data.[4] As the coverage of this dataset is limited to countries part of the LRTAP-framework, with consecutive data only from 1985, it is not well suited for constructing synthetic controls. Instead, I apply a newly developed database on SO_2, NO_x and VOC emissions for all countries in the world for the time period 1970-2008 (JRC, 2012).[5] Combining the synthetic control method (SCM in the following) with the global database allows me to address several shortcomings in the previous literature.

First, as participation in international protocols is voluntary, there is likely to be a self-selection bias. In particular, countries that are already on a downward-sloping path might be more inclined to join. If this is the case, it would lead to a violation of the key identifying

[3] See Section 3 for a more comprehensive overview of findings in the previous literature.
[4] An exception is Aakvik and Tjøtta (2011), see Section 3.
[5] The emissions database is constructed by using internationally reported activity data and assumptions on activity-specific emissions factors. See Section 5 for more details.

assumption underlying any potential outcomes framework: the *common trends* assumption. To tackle this problem, I construct a unique synthetic control unit for each treated country that mimics the pre-treatment trend in emissions and important drivers as closely as possible. This ensures that the estimated treatment effects are conditional on a similar pre-treatment trend.[6] By using the global database, I can construct a larger pool of potential control countries than previously possible, improving the chances of finding a good pre-treatment match.

Second, to recover unbiased estimates, there cannot be *spillovers* to the control group – an assumption that is hard to meet in the case of large scale interventions like multilateral agreements. The first two assumptions also constitute an inherent trade-off as potential control countries that are similar to the treated country, and hence more likely to meet the common trends assumption, may at the same time be more likely to be (indirectly) affected by the intervention. Geographical and political proximity will likely facilitate diffusion of new policies and technological solutions, and if nearby countries are used as controls, it could potentially lead to an underestimation of the treatment effect.[7] Further, as certain abatement measures are complementary across pollutants, like switching fuels or enhancing energy efficiency, a protocol targeting SO_2 could also have an effect on NO_x emissions, and vice versa. If such complementaries are substantial, it could further underestimate effects of international cooperation if countries in the control group have ratified other protocols within the LRTAP framework. As previous studies rely on a sample of (mainly European) countries that signed the 1979 LRTAP Convention, it might downward-bias treatment effects if favorable spillovers are large. Here, I aim to mitigate such concerns by expanding the sample to non-LRTAP countries, which allows me to run robustness checks where I exclude countries that are likely to be indirectly affected by a specific LRTAP protocol, such as non-ratifying countries in close geographical proximity.[8]

[6] Other potential approaches are to combine a difference-in-difference (DiD) with matching on lagged outcome variables, or by combining a DiD with country-specific linear or quadratic time trends. The latter approach will likely absorb large parts of the treatment effect, as treatment is redefined as deviations from the imposed trend. The SCM, by contrast, imposes no such restrictions. Previous literature has conducted a systematic comparison of the three approaches, and found the SCM to be the least biased estimator (Powell, 2017; O'Neill et al., 2016).

[7] Treatment effects could also be overestimated if negative spillovers, such as emissions leakage, dominate.

[8] In the main estimation, I keep countries that have both signed and ratified a specific protocol (e.g., the Helsinki protocol) in the treatment group. The donor pool consists of a trimmed sample of non-ratifying countries, where the criteria for trimming the donor pool are described in Section 5.2. For example, low-income countries are excluded. I also remove countries that have signed but not ratified the protocol in question, as it is not clear how these should be treated. This only applies to a few countries, and only the Geneva protocol. I run several robustness checks to examine the sensitivity of the main results to making changes to the donor pool, such as removing all LRTAP-

Third, while previous studies tend to use ratification or entry into force as the "intervention" date, countries may start reducing emissions before the formal implementation of the protocol due to rational expectations, or as a consequence of the dialog leading up to ratification. If there are signs of anticipation, Abadie (2012) suggests to backdate the intervention to a period before any anticipation effect can be expected in order to capture the full extent of the treatment effect. For the Helsinki protocol on SO_2, the choice of intervention date is particularity challenging as the focus of the 1979 LRTAP Convention was to combat SO_2 emissions. We might therefore expect to see effects materializing before the formal Helsinki protocol meeting in 1985. Additionally, the baseline year in the Helsinki protocol was set to 1980, and if this was known in advance, countries had an incentive to cut emissions in the years leading up to the meeting. To address potential anticipation effects, I define the intervention date of all protocols to the baseline year in the respective protocols. For the Sofia and Geneva protocols, this corresponds to the year before the protocol meetings.

Fourth, massive structural changes took place in Eastern Europe in the period analyzed, such as the fall of the Soviet Union and the reunification of Germany, potentially confounding the estimated treatment effects. In contrast to previous studies, I exclude all countries heavily affected by the collapse of the Soviet Union, such as former USSR-countries, former Yugoslavia, and Germany.

Fifth, average effects might conceal substantial heterogeneity. By applying the SCM, I can estimate country-specific treatment effects and thereby unveil which countries increased or decreased emissions compared to the constructed counterfactual. To summarize results, I pool country estimates to arrive at an average, protocol-specific treatment effect.[9] To evaluate the statistical significance of the pooled estimate, I compare the mean of the percentile ranks of the effects of the treated countries to those of donor countries. Since the (mean of) percentile ranks has a known distribution under the null hypothesis, I am able to perform exact inference. By

countries that did not ratify the protocol in question. While most LRTAP countries are high-income, there are still several high- and medium-income countries outside the framework that could potentially be used to construct synthetic controls, such as Australia, New Zealand, and Japan.

[9] As described in Section 5.2, I normalize emissions per capita by setting emissions of the affected country to be equal to 100 in the year before the intervention. A similar approach is taken in Cavallo et al. (2013) and Almer and Winkler (2017). Normalizing emissions eases comparability of effects across countries, and also lets me pool estimates to arrive at an average (unweighted) treatment effect.

inverting the mean rank statistic, I can construct confidence intervals for the pooled estimate.[10] These confidence intervals are also used to calculate an alternative point estimate (the Hodges-Lehman estimate), by taking the mean of the upper and lower confidence bounds.

While the aim of this paper is to improve on past estimates on the effects of pollution protocols, by carefully considering and addressing key identifying assumptions, it is worth reminding the reader that establishing causal inference of large scale interventions is inherently difficult. In an increasingly globalized world, however, international agreements are bound to play an important role also in the future. Applying the best available tools and data might be our best option if we wish to shed light on the effectiveness of international environmental agreements.

Results from the empirical examination show that all three LRTAP protocols induced emissions reductions beyond a (synthetic) counterfactual development. Using 1980 as the intervention year for the Helsinki protocol, I find that emissions were 23% lower than the synthetic control five years into the treatment period, and 22% lower after ten years.[11] The deviation from the control group hence occurred in the first five years. The large treatment effect of the Helsinki protocol contrasts the null finding in most previous studies. After disentangling potential causes of this discrepancy, I find the way non-parallel trends are dealt with to be the most important explanation (see below). Examining the Sofia protocol using 1987 as the intervention year, I find that emissions were 11% lower than the counterfactual after five years, which is comparable to previous findings.[12] After ten years, the corresponding estimate is 18%.[13] For the Geneva protocol, treatment effects after five and ten years are 15% and 20%, respectively.[14] Using the rank-based inference procedure, I find that the pooled treatment effects of each of the three protocols are statistically significant at a 1% level. The synthetic control units mimic the pre-treatment development in emissions relatively closely, and estimated treatment effects are robust to several adjustments to the predictor set and donor pool.

The empirical examination sheds light on two important methodological issues. First,

[10] The procedures to conduct inference on the pooled estimate and construct confidence intervals are similar to procedures described in Dube and Zipperer (2015) and Gobillon and Magnac (2016). The procedure of using mean percentile ranks to evaluate statistical significance of the pooled estimate can be seen as an extension of the single event, placebo-based inference used in Abadie et al. (2010).

[11] For the Hodges-Lehman (HL) estimate, the effects are -20% after five years, and -18 % after ten years.

[12] Bratberg et al. (2005) estimate an average, annual treatment effect of around -2.1%.

[13] The Hodges-Lehman (HL) estimates are roughly of the same magnitude.

[14] The corresponding Hodges-Lehman (HL) estimates are -19% after five years and -17 % after ten years.

changing the intervention year to the year protocols entered intro force, lowers the treatment effects significantly. This is particularly the case for the Helsinki protocol on SO_2, where changing the start date to the year the Helsinki protocol entered into force (1987) renders a small and insignificant treatment effect. This highlights the importance of accounting for anticipation effects to capture the full extent of the treatment.[15]

Second, the empirical investigation illustrates an important shortcoming of traditional ways of dealing with non-parallel trends, such as combining a difference-in-difference (DiD) with country-specific time trends. By applying the DiD setup from one of the recent studies on the Helsinki protocol (Aakvik and Tjøtta, 2011), I find that their choice of treatment date (1986) and control group (LRTAP countries only) explain some of the discrepancy, but the main reason for their small and insignificant treatment effect is due to the inclusion of linear or quadratic country-specific time trends. While their motivation for including such trends is to address violation of the parallel trends assumption, the imposed trends seem to absorb most of the treatment effect. The synthetic control method offers an alternative way of controlling for different pre-intervention trends that avoids the risk of absorbing treatment effects, and can be seen as an extension of the DiD framework to account for time-varying confounders.

Overall, the results in this paper suggest that international agreements have been successful in reducing emissions beyond what they would have been in absence of the interventions. This finding contrasts the pessimistic predictions from the theoretical literature that free-riding incentives will render such agreements ineffective. The results also highlight important methodological issues, such as accounting for non-parallel trends in a suitable manner, and to carefully define the treatment window. The findings also help explain why previous studies using a DiD strategy and officially reported data tend to find small or no effects of the LRTAP protocols.

The remainder of the paper is structured as follows. Section 2 gives the historical background of the different protocols. Section 3 reviews the previous literature evaluating the LRTAP protocols. Section 4 presents the methodology, while Section 5 describes the data. Section 6 presents the results, and Section 7 concludes.

[15] An alternative interpretation is that countries experiencing a decline in emissions were more inclined to ratify the protocol. However, given we are interested in estimating the effect of the LRTAP framework, the natural intervention date would be the time of the first Convention.

2 Background

2.1 The Convention on Long-range Transboundary Air Pollution

In the 1960s scientists started to unravel the link between sulfur emissions (SO_2) in continental Europe and the acidification of Scandinavian lakes. While the environmental damages were first noted in the early 1920s, the idea that air pollutants could travel thousands of kilometers before depositing and creating damage to lakes, rivers and forest didn't receive notable attention until the 1960s. The 1972 United Nations Conference on the Human Environment in Stockholm signaled the start of an international initiative to combat transboundry pollution. While several countries remained skeptical of the proclaimed relationship between transboundary pollution and the environmental damages in Scandinavia, new studies in the period 1972-1977 confirmed the hypothesis, which led to a broader scientific consensus (UNECE, 2015).

Having recognized the severity of the problem, and thereby the need for international cooperation, a high-level meeting of the UN Economic Commission for Europe on the Protection of the Environment was held in November 1979 in Geneva. The meeting is formally known as *the Convention on Long-range Transboundary Air Pollution* (LRTAP). Article two of the LRTAP Convention states that "The Contracting Parties (...) shall endeavour to limit and, as far as possible, gradually reduce and prevent air pollution including long-range transboundary air pollution."[16] The 1979 Convention was largely a framework agreement, formulating general principles for cooperation on air pollution abatement. It has later been extended by eight specific protocols containing legally binding targets for emission reductions.[17] Six of these protocols targeted SO_2, NO_x or VOC emissions, and are listed in Table 1.[18]

The 1984 *Geneva Protocol on Long-term Financing of the Cooperative Programme for Monitoring and Evaluation of the Long-range Transmission of Air Pollutants in Europe* (EMEP) was the first protocol to be signed as part of the LRTAP framework. The protocol did not set any emission reduction targets, but provided a financing scheme to fund future activities and

[16] The Convention text is available here: `http://www.unece.org/fileadmin/DAM/env/lrtap/full%20text/1979.CLRTAP.e.pdf`.

[17] A convention is a formal agreement between states, and is synonymous with the term treaty. The term protocol is used for an additional legal instrument that complements and adds to a treaty. A protocol is optional because it is not automatically binding for States that have ratified the initial treaty; States must independently ratify a protocol.

[18] For more information on the different pollutants, and how they are linked to each other, see Appendix A.2.

Table 1: International conventions and protocols part of the LRTAP framework

Short name	**Category**	**Pollutant(s)**	**Open for signature**	**Entry into force**	**Baseline year(s)**
LRTAP	Convention		Nov 1979	Mar 1983	
EMEP	Protocol		Sep 1984	Jan 1988	
Helsinki	Protocol	SO_2	Jul 1985	Sep 1987	1980
Sofia	Protocol	NO_X	Oct 1988	Feb 1991	1987
Geneva	Protocol	VOCs	Nov 1991	Sep 1997	1984-1990
Oslo	Protocol	SO_2	Jun 1994	Aug 1998	
Gothenburg	Protocol	SO_2, NO_X, VOC	Nov 1999	May 2005	

Notes: In addition to the six protocols listed, the LRTAP framework also includes two protocols addressing persistent organic pollutants (POPs) and heavy metals: the 1998 *Aarhus Protocol on Persistent Organic Pollutants (POPs)* and the 1998 *Aarhus Protocol on Heavy Metals*.

provide information on emissions, transport, and deposition of air pollution. In that way the protocol represented the backbone of the Convention.

2.2 The Helsinki, Sofia and Geneva protocols

The first protocol to contain specific emission reduction targets was the 1985 *Helsinki Protocol on the Reduction of Sulphur Emissions or their Transboundary Fluxes by at least 30 per cent* (the Helsinki protocol in the following).[19] SO_2 emissions had already been established as an important source of acidification of rivers and lakes, and was therefore a natural starting point for the first international protocol. The Helsinki protocol opened for signature in July 1985, and entered into force in September 1987. The protocol committed ratifiers to reduce SO_2 emissions by at least 30% compared to 1980 levels, as soon as possible or by 1993.

As more scientific evidence was provided, it became clear that other pollutants, like nitrogen oxides, were also contributing to acidification, and had to be addressed within the international framework. This led to the 1988 *Sofia Protocol concerning the Control of Emissions of Nitrogen Oxides or their Transboundary Fluxes* (the Sofia protocol in the following).[20] The protocol required countries to introduce pollution control measures for the largest existing stationary sources, and to apply national emission standards to major new stationary and mobile sources.

[19]The protocol text is available here: `http://www.unece.org/fileadmin/DAM/env/documents/2012/EB/1985.Sulphur.e.pdf`

[20]The protocol text is available here: `http://www.unece.org/fileadmin/DAM/env/lrtap/full%20text/1988.NOX.e.pdf`.

The aim stated in the protocol was to reduce NO_x emissions to 1987 levels by December 1994.[21]

In subsequent years, countries recognized that volatile organic compounds (VOCs), in addition to NO_x, were contributing to the formation of ground-level ozone and other photochemical oxidant products, causing damage to vegetation and crops. To reduce VOCs, countries adopted the *1991 Geneva Protocol concerning the Control of Emissions of Volatile Organic Compounds or their Transboundary Fluxes* (the Geneva protocol in the following).[22] Under the Geneva protocol, countries had the opportunity to choose between three different emission reduction targets: a 30 % reduction by 1999 (using a year between 1984 and 1990 as the benchmark)[23], a 30 % reduction by 1999 within a so-called Tropospheric Ozone Management Area and ensuring that 1999 emissions did not exceed 1988 levels[24], or a stabilization of emission by 1999 at the same levels as in 1988 - given the 1988 levels did not exceed a specified threshold.[25]

2.3 The Oslo and Gothenburg protocols

The Helsinki protocol was replaced by the *1994 Oslo Protocol on Further Reduction of Sulphur Emissions* (the Oslo protocol in the following). While previous protocols roughly prescribed the same percentage emission reductions for all countries, the Oslo protocol derived required emission reductions from cost-effectiveness and effect-based principles.[26] The Oslo, Sofia, and Geneva protocols were later replaced by one single protocol: the 1999 *Gothenburg Protocol to Abate Acidification, Eutrophication and Ground-level Ozone* (the Gothenburg Protocol in the following). The protocol was the first multi-pollutant protocol, covering four different pollutants; SO_2, NO_x, ammonia (NH3), and VOCs. Similar to the Oslo protocol, the Gothenburg protocol used the principle of cost-effectiveness to set national emission caps.

[21] The reference year was 1987 for all countries except the United States, which used 1978 as the reference year.

[22] The protocol text is available here: `http://www.unece.org/fileadmin/DAM/env/lrtap/full%20text/1991.VOC.e.pdf`.

[23] This option was chosen by Austria, Belgium, Estonia, Finland, France, Germany, Netherlands, Portugal, Spain, Sweden, and the United Kingdom (with 1988 as base year), by Denmark (with 1985 as base year), by Liechtenstein, Switzerland and the United States (with 1984 as base year), and by Czech Republic, Italy, Luxembourg, Monaco and Slovakia (with 1990 as base year). Source: `https://www.unece.org/fileadmin/DAM//env/lrtap/vola_h1.htm`

[24] This option was chosen by Norway (with 1989 as the benchmark year) and Canada (with 1988 as the benchmark year). See Annex I to the Protocol for a definition of a Tropospheric Ozone Management Area.

[25] This option was chosen by Bulgaria, Greece and Hungary.

[26] Specifically, each country's required emission reductions were based on the results of a modeled relationship between SO_2 emissions and the exposure of different ecosystems.

Table 2: Ratification year of LRTAP conventions and protocols, by country

Country name	**LRTAP**	**Helsinki (SO_2)**	**Sofia (NO_X)**	**Geneva (VOC)**	**Oslo (SO_2)**	**Gothenburg (SO_2,NO_X,VOC)**
Austria	1982	1987	1990	1994	1998	
Belgium	1982	1989	2000	2000	2000	2007
Canada	1981	1985	1991		1997	
Cyprus	1991		2004		2006	2007
Denmark	1982	1986	1993	1996	1997	2002
Finland	1981	1986	1990	1994	1998	2003
France	1981	1986	1989	1997	1997	2007
Greece	1983		1998		1998	
Iceland	1983					
Ireland	1982		1994		1998	
Italy	1982	1990	1992	1995	1998	
Luxembourg	1982	1987	1990	1993	1996	2001
Malta	1997					
Netherlands	1982	1986	1989	1993	1995	2004
Norway	1981	1986	1989	1993	1995	2002
Portugal	1980					2005
Romania	1991					2003
Spain	1982		1990	1994	1997	2005
Sweden	1981	1986	1990	1993	1995	2002
Switzerland	1983	1987	1990	1994	1998	2005
Turkey	1983					
United Kingdom	1982		1990	1994	1996	2005
United States	1981		1989			2004

Notes: Table shows countries that have ratified the LRTAP Convention (before 2000), and that are included in the main sample in the analysis. The years indicate the country-specific ratification year of each protocol. Several countries are excluded from the sample based on large structural changes taking place in the period analyzed, such as former USSR countries, former Yugoslavia (incl. Albania), former Czechoslovakia, Bulgaria, Germany, and Poland. Small islands and microstates like Monaco are also excluded. Some countries lack data on pollution and/or GDP, and therefore need to be excluded from the analysis. There are five countries that have signed but not ratified the Geneva protocol: Canada, Greece, Portugal, Ukraine, and the United States. See Table A.1 in the Appendix for a complete list of ratifying countries, and Section 5.2 for a complete list of the exclusion criteria.

2.4 Ratification of the protocols

Table 2 lists LRTAP countries included in the analysis in Section 6.[27] While Belgium, Denmark, Finland, France, Luxembourg, the Netherlands, Norway, Sweden and Switzerland ratified all five subsequent protocols on SO_2, NO_x and VOCs, the rest of the countries ratified four or less. Iceland, Turkey and Malta have to date only ratified the initial LRTAP Convention, while Portugal and Romania have only ratified the Gothenburg protocol. Five countries have signed but not ratified the Geneva protocol (Canada, Greece, Portugal, Ukraine, and the United States).

[27]Table A.1 in the Appendix gives a complete list of countries ratifying the 1979 LRTAP Convention.

3 The effects of the LRTAP protocols: previous findings

Over the past decades, several studies have emerged to shed light on the effectiveness of the different LRTAP protocols.[28] In particular, the 1985 Helsinki protocol has been subject to several empirical evaluations. In an early study, Murdoch et al. (1997) investigate the effects of the 1985 Helsinki and 1988 Sofia Protocols. Using a spatial lag model with data for 25 European countries over the time period 1980-1990, the authors find that the Helsinki protocol has been more effective in reducing emissions than the Sofia protocol.[29] While they cautiously conclude that SO_2 emissions have been easer to combat than NO_x, their study only includes countries that were covered by the protocols. Their findings hence do not constitute evidence on how emissions would have evolved in absence of treatment.

In a subsequent study, Murdoch et al. (2003) focus on SO_2 emissions, and use a joint spatial probit and spatial lag equation to estimate both the participation decision and the level of participation in the Helsinki protocol. Using the same dataset as in Murdoch et al. (1997), they find that voluntary cutbacks beyond the emission target gives incentives to free ride. Again, the study does not say anything about the counterfactual, but focuses on the strategic interaction among ratifiers of the protocol. In a closely related study, Finus and Tjøtta (2003) use a numerical model to test if countries ratifying the 1994 Oslo protocol reduced SO_2 emissions beyond the numerical calibrated Nash equilibrium. Comparing actual reductions to a simulated Nash equilibrium, they find that the targets for the Oslo protocol are very close to the simulated Nash equilibrium, and the protocol hence provided little emission cuts beyond Nash behavior.

Focusing on NO_x, Bratberg et al. (2005) estimate the effects of the 1988 Sofia protocol using a differences-in-differences (DiD) approach. They use a sample of 23 European countries for the period 1985-1996 to evaluate the effect, and find evidence that the protocol led to emission reductions slightly greater than what they would have been in absence of the protocol. The yearly reductions in emissions are found to be around 2.1% greater for countries ratifying the Sofia protocol compared to non-ratifiers. In a similar type of set-up, Ringquist and Kostadinova (2005) estimate the effect of the 1985 Helsinki Protocol. Using data on emissions for 19 Eu-

[28] For an overview of empirical studies, see e.g., Houghton and Naughton (2016).

[29] The authors suggest that the stationary sources of SO_2 emissions, together with the substance traveling shorter distances, makes SO_2 somewhat easer to control than NO_x emissions.

ropean countries for the time period 1980-1994, the authors find that while countries ratifying the Helsinki Protocol experienced significant emission reductions, the protocol itself had no significant effect on emissions. The same conclusion in reached in Naughton (2010). Using a sample of 16 European countries for the time period 1980-2000, Naughton (2010) estimates the effects of the Helsinki, Oslo and Sofia protocols. Applying a 2SLS spatial lag model, the author finds no evidence of an effect of the two first protocols, while the Sofia protocol reduced NO_x emission levels and trend on average.

A common feature of the previous studies on the LRTAP protocols is the use of a small sample consisting of only European countries, as well as the use of a short pre-intervention time period. Aakvik and Tjøtta (2011) take the literature a step forward by exploiting a newly assembled dataset on SO_2 emissions dating back to 1960, and covering in total 30 European countries. Using a DiD approach, they estimate the effect of the 1985 Helsinki and 1994 Oslo Protocols. Controlling for country-specific linear and quadratic time trends, and using 1986-1993 as the treatment window for the Helsinki protocol and 1995-2001 for the Oslo protocol, the authors find no significant effects of the protocols. While the study addresses some of the limitations of the previous literature by applying a dataset covering more countries over a longer time period, their study also has limitations. First, including country trends may absorb parts of the treatment effect, as the treatment effect is now measured as deviations from a linear or quadratic trend. Second, as the case with previous studies, only including European countries might introduce a downward-bias due to policy and technology spillovers. Lastly, by using 1986 as the intervention year, the analysis does not account for potential anticipation effects.

Compared to previous studies, I contribute to the literature in at least four aspects. First, I apply a relatively new methodology for evaluating potential effects of the international protocols that addresses the concern of different pre-treatment trends. By combining the synthetic control method with a newly constructed database, which dates back to 1970, I am able to take into account different pre-treatment trends. Specifically, I construct a unique synthetic counterfactual development for each individual country. Second, as the new dataset covers a large number of countries, I am better equipped to address problems of spillovers. By excluding nearby, non-ratifying countries from the control group, and bringing in countries outside Europe and North America, I am able to mitigate problems of spillovers and complementaries between protocols. Third, by constructing a synthetic control group for each treated country, I can investigate country-specific treatment effect. Fourth, using the new data source, I take into account potential anticipation effects by backdating the treatment date for the Helsinki protocol to the time of the first Convention meeting.

4 Methodology

In this paper, I set out to estimate causal effects of the LRTAP protocols on emissions of SO_2, NO_x and VOCs. To address the problems of different pre-treatment trends, I apply the synthetic control method, which uses a weighted combination of control countries to construct "synthetic" counterfactual. The method was first introduced by Abadie and Gardeazabal (2003), and later extended in Abadie et al. (2010), where they estimate the effect of a large tobacco control program in California.[30] I also draw on Dube and Zipperer (2015) and Gobillon and Magnac (2016) when conducting inference on the pooled estimate and constructing confidence intervals.

4.1 The synthetic control estimator

4.1.1 A single treated unit

I start by presenting a framework for the case of a single treated country. Assume that we have data for a sample of $J+1$ countries, where $j=1$ denotes the "treated" country, i.e., the country affected by the policy intervention, and $j=2,\ldots,J+1$ are countries unaffected by the intervention, i.e., the "donor pool". In our setting, the intervention is participation in an international pollution protocol, and the outcome of interest is emissions of SO_2, NO_x or VOCs. Further, assume that the data spans T periods, where T_0 is the period prior to intervention. Denoting the intervention as D, the synthetic control approach assumes that the observed outcome, Y_{jt}, is the

[30] The method has later been applied to a wide range of topics. Examples include the economic impact of natural resource endowment (Mideksa, 2013), the effect of economic liberalization on GDP (Billmeier and Nannicini, 2013), impact of catastrophic natural disasters on economic growth (Cavallo et al., 2013), the effects of the German reunification on economic costs (Abadie et al., 2015), the economic costs of organized crime (Pinotti, 2015) and the effects of the Kyoto protocol on CO_2 emissions (Almer and Winkler, 2017).

effect from the treatment, $\alpha_{jt}D_{jt}$, and the counterfactual outcome, Y_{jt}^N:

$$Y_{jt} = \alpha_{jt}D_{jt} + Y_{jt}^N = \alpha_{jt}D_{jt} + \theta_t Z_j + \lambda_t \mu_j + \delta_t + \varepsilon_{jt} \tag{1}$$

Here Z_j is a vector of observed covariates not affected by the intervention, θ_t is a vector of unknown parameters, δ_t is a common time factor and ε_{jt} is the idiosyncratic error term. In a standard difference-in-differences (DiD) framework, both Z_j and δ_t can be accounted for by comparing the difference between the treatment group and the control group before and after the intervention. As long as the covariates Z_j do not vary over time, and the time trend δ_t is common to all countries, the terms will be differenced out in a DiD set-up. What is left, however, is the term $\lambda_t \mu_j$. Here λ_t is a vector of unobserved *time-varying* factors and μ_j are the unknown factor loadings. If the factor loadings differ across countries, the assumption of parallel trends for the treated and control countries in absence of intervention will likely be violated. However, if we knew the true factor loadings μ_1 for the treated country, we could construct an unbiased control by using donor states whose factor loadings average to μ_1

The idea of the synthetic control method is to construct a vector of weights **W** over J donor states such that the weighted combination of donor states closely mimics the outcome of the treated country in the pre-intervention period. This weighted combination of donor units is called the synthetic control. Given a good match, we can difference out the time-varying term $\lambda_t \mu_j$. More formally, for the treated country, I define the $(k \times 1)$ vector of pre-treatment characteristics as $X_1 = (Z_1', Y_j^{K_1}, ..., Y_j^{K_L})$, where $Y_j^{K_i}$ are L linear combinations of pre-treatment outcomes. Analogously, I define the $(k \times J)$ matrix containing the same characteristics for the J donor countries as X_0. The synthetic control procedure chooses donor weights **W** to minimize the distance between pre-treatment characteristics X_1 and X_0 of the treated country and untreated countries. More specifically, the method minimizes the mean square prediction error (MSPE) over k pre-treatment characteristics:

$$\sum_{m=1}^{k} v_m (X_{1m} - X_{0m}W)^2, \tag{2}$$

where v_m measures the relative importance of the m^{th} predictor. Given the optimal weights w_j^*

for each of the $j = 2,\ldots,N$ donors, the synthetic control at any time t is simply the weighted combination of the outcome variable (i.e., pollution) in the donor countries: $\sum_{j=2}^{N} w_j^* Y_{jt}$.[31] The estimate of the treatment effect α_{1t} is therefore the difference between pollution in the treated country Y_{1t} and pollution in the synthetic country $\sum_j w_j^* Y_{jt}$ at any post-treatment time $t \geq T_0$:

$$\hat{\alpha}_{1t} = Y_{1t} - \sum_{j=2}^{N} w_j^* Y_{jt} \tag{3}$$

In the post-intervention period $t = T_0,\ldots,T$, the average difference between the treated and synthetic control outcomes is given by

$$\hat{\beta}_1 = \frac{1}{T} \sum_{t=T_0}^{T} (Y_{1t} - \sum_{j=2}^{N} w_j^* Y_{jt}) \tag{4}$$

In the analysis, the outcome variable is normalized to 100 in the year prior to treatment (see Section 5.2). This means that we can interpret $\hat{\beta}_1$ as the average difference in percentage points between the treated and the synthetic counterfactual development.

4.1.2 Multiple treated units: pooled estimate

In the case of the LRTAP protocols, there are multiple treated countries. I therefore generalize the framework described above to multiple units. Denoting the treated countries by subscript e, where $e = 1,\ldots,E$, I calculate an annual, country-specific treatment effect $\hat{\alpha}_{e1t}$ and an average, country-specific treatment effect $\hat{\beta}_{e1}$ by using equations 3 and 4. The average *pooled* treatment effect can be expressed as:

$$\bar{\alpha}_{e1t} = \frac{1}{E} \sum_{e=1}^{E} \hat{\alpha}_{e1t} \quad \bar{\beta}_{e1} = \frac{1}{E} \sum_{e=1}^{E} \hat{\beta}_{e1}, \tag{5}$$

where $\bar{\alpha}_{e1t}$ is the pooled treatment effect for a given year, and $\bar{\beta}_{e1}$ is the pooled treatment effect averaged over the post-treatment period. I also calculate an alternative pooled treatment effect, the Hodges-Lehman (HL) pooled estimate, which I explain in detail in Section 4.4.

[31]The weights are restricted to sum to one. This implies that synthetic controls are weighted averages of the units in the donor pool. Restricting country weights to sum to one may be warranted only if the dependent variable is rescaled, so it is not affected by country size. As described in Section 5.2, I use normalized variables of pollution per capita as the outcome variable, which would warrant such a restriction.

4.2 Statistical inference

4.2.1 A single treated unit

To assess the statistical significance of a single country's estimated treatment effect, I use placebo-based inference. This involves running a number of falsification tests, or "placebo tests", for the countries in the donor pool. The estimated treatment effect for the treated unit is then compared to the distribution of placebo effects. Specifically, I estimate treatment effects $\hat{\beta}_j$ for each of the $j = 2, ..., N$ donor countries by repeating the procedure described in Section 4.1, but using the remaining $N-2$ donor counties. These placebo runs are used to evaluate the statistical significance of the true treatment estimate. In the case of a single treated country, I compare the magnitude of the treatment effect for the treated county to the treatment effects of the placebo runs.[32] I then rank the treatment effects according to magnitude. This allows me to construct a percentile rank statistics p for the treated country:

$$p_{1t} = \hat{F}(\alpha_{1t}) \quad p_1 = \hat{F}(\beta_1), \tag{6}$$

where F is the empirical cumulative distribution function (CDF) of the coefficients $\hat{\alpha}_{jt}$ or $\hat{\beta}_j$. As the percentile rank is approximately uniformly distributed, I can determine whether the rank of the treated state, p_1, lies in the tails of the distribution. Using a two-sided statistical significance level of 5 percent, I reject the null of $\beta_1 = 0$ when $p_1 < 0.025$ or $p_1 > 0.975$.[33]

4.2.2 Multiple treated units: pooled estimate

To conduct inference on the pooled treatment effect in equation 5, I construct a test statistic $\bar{p}$ which is the mean of the percentile ranks of treated countries:

$$\bar{p}_t = \frac{1}{E}\sum_{e=1}^{E} p_{et} \quad \bar{p} = \frac{1}{E}\sum_{e=1}^{E} p_e \tag{7}$$

[32] To account for the fact that a poor pre-treatment fit might give rise to larger post-treatment deviations, I trim the donor pool based on pre-treatment fit. Specifically, I trim the donor pool down to the 42 countries with the lowest mean square prediction error (MSPE).

[33] Note that the number of available donors limits the range of confidence levels I can implement for a single treated event. In order to asses a two-sided 5 percent level of significance, I need at least 41 donor countries.

If we assume that ranks are independent across treated countries, the exact distribution of $\bar{p}$ can be calculated using the Irwin-Hall distribution of the sum of E independent uniform random variables. The procedure is described in detail in Appendices B.1 - B.2. Alternatively, we can form a distribution of the mean percentile ranks by randomly permuting the treatment status, see Appendix B.3.[34] The permutation exercise is far more computationally intensive than using the Irwin-Hall distribution. Also, the small number of actually observed percentile ranks will influence the cut-off values. I therefore focus on the cut-off values from the Irwin-Hall distribution when evaluating statistical significance (see Appendix Table B.3), but use the cut-off values from the permutation procedure in robustness checks.[35]

4.3 Constructing confidence intervals

4.3.1 A single treated unit

In the case of a single treated country, we can invert the percentile ranks, p_{1t}, to construct confidence sets. Inverting the percentile rank means that I ask for what values of τ does the following inequality hold:

$$0.025 \geq \hat{F}_{1t}(\alpha_{1t} - \tau) \geq 0.975 \tag{8}$$

The term $\hat{F}_{1t}(\alpha_{1t} - \tau)$ is referred to as the *adjusted* country-specific rank, $p_{1t}(\tau)$. The 95 percent confidence interval is the set of τ not rejected using the critical values 0.025 and 0.0975.

4.3.2 Multiple treated units: pooled estimate

To construct confidence intervals for the pooled effect, I invert the mean rank statistic $\bar{p}_t$. This means that I ask for what values of τ does the following inequality hold:

$$\textit{Lower critical value} \geq \frac{1}{E}\sum_{e=1}^{E} \hat{F}_{e1t}(\alpha_{e1t} - \tau) \geq \textit{Upper critical value}, \tag{9}$$

[34]The permutation procedure has similarities to the procedures described in Section 4.5 in Dube and Zipperer (2015) and in the Results section in Gobillon and Magnac (2016).

[35]Additionally, I also address the potential problem of rank dependency by performing a randomization procedure that constrains the permutation of treatment status by forcing the "treated" countries to be located geographically close to each other. The procedure is described in detail in Appendix B.4, while robustness checks with these alternative cut-off values are presented in Section 6.5.

where $\frac{1}{E}\sum_{e=1}^{E}\hat{F}_{e1t}(\alpha_{e1t}-\tau)$ is the *mean adjusted rank*, $\bar{p}_t(\tau)$.[36] The 95 percent confidence interval for the pooled effect is the set of τ such that the mean adjusted rank $\bar{p}_t(\tau)$ lies within the critical values presented in Appendix Table B.3.

4.4 The Hodges-Lehman (HL) point estimate

By collapsing the pooled confidence intervals, I get the Hodges-Lehman (HL) point estimate (Hodges Jr and Lehmann, 1963). The HL estimate is simply the mean of the upper and lower confidence bounds. In the case of a single treated country, the mean and the HL point estimate are the same. In the case of multiple treated countries, the mean and the HL point estimate are not necessarily the same. If outlying estimates of individual treatment effects heavily influence the mean estimate, the mean and the HL estimate will differ substantially. While the mean estimate has a more clear interpretation, the HL estimate is more robust to outliers.

4.5 Requirements and caveats

In the following, I describe the conditions that need to be in place for the synthetic control method to be an appropriate tool for evaluating a policy intervention.[37]

First, if the outcome variable is *highly volatile*, the synthetic control method may not be able to distinguish a treatment effect from random shocks to the outcome variable. In particular, if the magnitude of impacts from an intervention is similar to the volatility of the outcome variable, treatment effects are difficult to detect.[38]

Second, if potential control countries adopt a *similar type of intervention* as the one adopted by the treated country, they should not be included in the donor pool.[39] It is also important to eliminate from the donor pool any country that may have suffered *large idiosyncratic shocks* to the outcome of interest during the period analyzed.

Third, the *differences in the characteristics* of the affected country and the synthetic control

[36]A similar type of procedure is described in Gobillon and Magnac (2016).
[37]Several of the conditions are also relevant to other policy evaluation tools, including difference-in-differences.
[38]This problem arises if the volatility is intrinsic to the treated country. Common shocks affecting all other countries can be differentiated out by choosing a suitable synthetic control.
[39]As an example, Abadie et al. (2010) discard from the donor pool several states that adopted large-scale tobacco programs during the sample period of the study.

should not be too big. If a country had particularly low or particularly high levels of emissions before the treatment date relative to the countries in the donor pool, then no weighted average of countries in the donor pool will be able to closely reproduce the pre-intervention emissions for the country. As a way around this, Abadie (2012) suggests to transform the outcome to time differences or growth rates.[40]

Fourth, while countries in the donor pool should not be too different from the treated countries, they should at the same time be unaffected by the intervention. If *spillover effects* are likely to be substantial, it may be advisable to exclude countries expected to be indirectly affected. There is hence a tension between the issue of no spillovers and having comparable countries in the donor pool.

Fifth, the synthetic control estimator may be biased if forward looking countries *react in advance of the policy intervention*, or if certain components of the intervention are put in place before the formal implementation. If there are signs of anticipation, Abadie (2012) recommends to backdate the intervention to a period before any anticipation effect can be expected in order to capture the full extent of the treatment effect.

5 Data and descriptives

5.1 Data sources

Data used in the analysis is complied from several sources. Information on participation in environmental protocols is collected from the International Environmental Agreements database project (Mitchell, 2015).[41] The database contains information on when a protocol opened up for signature, the date it entered into force, as well as each country's signature and ratification date. Table 2 lists each country's ratification year for different protocols.

In order to apply the synthetic control method, I need emissions data for both affected and unaffected countries. Further, the dataset needs to span a pre and post intervention period. As the officially reported data to the European Monitoring and Evaluation Programme (EMEP)

[40] The same logic is used in a difference-in-differences framework; even if the level of the outcome variable cannot be reproduced, there are cases when a control group can reproduce the *changes* in the outcome variable for the treatment group.

[41] The data is available at `http://iea.uoregon.edu/`

only covers countries part of the LRTAP framework, and only dates back to 1980, I use a different source of data for the analysis: the Emission Database for Global Atmospheric Research (EDGAR in the following) (JRC, 2012).[42] The development of EDGAR is a joint project of the European Commission Joint Research Centre and the Netherlands Environmental Assessment Agency, and provides global emissions of air pollutants by country and sector.[43] The emissions data in EDGAR is derived by pairing internationally reported activity data with assumptions on sector- and technology-specific emissions factors. The bottom-up methodology is applied consistently for all world countries, resulting in country-sector specific emissions estimates on a wide range of pollutants, including SO_2, NO_X and VOCs. For each of the three pollutants, emissions data is available from 1970 to 2008.[44]

To evaluate the similarity of countries in the donor pool to treated countries, I collect data on the following country characteristics from the World Bank (The World Bank, 2015): GDP per capita (in constant 2005 US$), GDP growth, population growth, and the share of fossil fuels of total energy use. These variables can be used to exclude countries with very different pre-treatment characteristics, and can be used as predictors to construct the synthetic controls.[45] Note that even if pre-intervention emissions for the synthetic control and the treated country closely align, the synthetic control should also approximate the treated country in the values of the most important predictors of the outcome variable, such as GDP per capita.[46]

[42] The officially reported data to the EMEP is available at `http://www.ceip.at/ms/ceip_home1/ceip_home/webdab_emepdatabase/reported_emissiondata/`. The EMEP emissions data is available for the years 1980, 1985, and then annually from 1990 and onwards.

[43] The dataset is available at `edgar.jrc.ec.europa.eu/overview.php`. It was first made available in July 2010, but have been updated since. I use version 4.2 of the database, which was released on November 2011.

[44] Emissions factors are corrected for end-of-pipe abatement measures. For more details, see `http://edgar.jrc.ec.europa.eu/methodology.php`. Note that emissions in the EDGAR database may not necessarily correspond to the officially reported EMEP data, as the EDGAR database relies on a technology based emission factor approach. The same methodology is applied to all countries to ensure comparability. Note also that another dataset on SO_2 emissions (Stern, 2006), dating back to 1960, has been used in Finus and Tjøtta (2003). For consistency reasons, I use the same data source (EDGAR) for all three pollutants.

[45] Including additional predictors imply several trade-offs. First, many variables are missing for the pre-intervention period, or only available for a small sub-sample of countries. Including these variables as predictors will hence imply dropping a substantial number of countries from the analysis, or, alternatively, imply some form of imputation. Second, adding more predictions will necessarily lower the weights assigned to other predictors, like past emissions and GDP per capita, potentially leading to a poorer pre-intervention match on these variables.

[46] A relevant concept here is the so-called Environmental Kuznets Curve hypothesis, which postulates an inverted u-shaped relationship between pollution and GDP, see e.g., Dinda (2004). Although the empirical support for the hypothesis is mixed, comparing countries at different stages in the economic development could imply that richer countries are on a downward-sloping path while poorer countries are on an upward-sloping path. At the same time, the synthetic control method is designed to mitigate such problems, by constructing synthetic controls that approximate the development in emissions in the period before the intervention. Including GDP per capita as a

5.2 Defining the sample, treatment window and outcome variable

In the analysis, I focus on three interventions: the Helsinki protocol (on SO_2), the Sofia protocol (on NO_x) and the Geneva protocol (on VOCs). For each of the protocols, I start by defining treated units as countries signing and ratifying the protocol in question (before 2000), and the donor pool as countries not signing or ratifying the protocol. Based on the recommendations and caveats discussed in Section 4.5, I further restrict the sample, as well as define the treatment window, outcome variables and predictors. The adjustments are described in detail below, while a summary of the restriction criteria are presented in Appendix Table C.2.

5.2.1 Restricting the treatment group and donor pool

Initially, the emissions dataset covers over 170 countries. However, I make several adjustments that substantially lowers the number of countries in the sample. First, countries should not experience country-specific structural shocks to the outcome variable that coincide with the intervention. Based on this, I exclude countries heavily affected by the fall of the Soviet Union, such as former USSR countries, former Yugoslavia, former Czechoslovakia, Germany, and Poland, and countries experiencing long-lasting conflicts and wars during the treatment period. Next, as the majority of treated countries are high-income countries, I exclude the poorest quintile from the sample.[47] Lastly, I exclude small island states and microstates, such as Monaco and Lichtenstein, as well as countries with an extremely volatile development in emissions.[48]

A more difficult question is how to deal with spillovers. Spillovers, or indirect effects, can both increase or decrease emissions in countries not covered by the protocol in question. At the one hand, technology and policy diffusion might lead countries not covered by the protocol to reduce their emissions. This might particularly be the case for similar countries in close geographical proximity to treated countries. Also, European countries not covered by the protocol might be affected via new EU directives triggered by the protocol. If this is the case, including

predictor also ensures that income levels are not too different.

[47] Specifically, I exclude all countries with GDP per capita in the lowest 20th percentile in 1980.

[48] I use a criteria where I exclude countries if emissions in the peak year is more than three times higher than the minimum emissions in the period analyzed. I also exclude Norway from the treatment group when analyzing the Geneva protocol due to the drastic fluctuations in VOCs caused by the accelerating oil production from 1975. While storage and transportation of crude oil have large impacts on VOCs due to evaporation of chemicals, NO_x and SO_2 are primarily caused by fuel combustion, see Section A.2 in the Appendix.

these countries in the donor pool will likely underestimate the treatment effect. The same is true if there are strong complementarities between abatement measures for the three pollutants, and countries in the donor pool have ratified another LRTAP-protocol. In particular, SO_2 and NO_x are often emitted as co-pollutants, and efforts to reduce one of these pollutants, like switching fuel from coal to gas or enhancing energy efficiency, will likely effect both pollutants.[49] On the other hand, spillovers can also take the form of emission leakage through re-allocation of pollution-intensive industries, or via input markets, which could increase emissions in countries not covered by the protocol. This would instead overestimate the effect of the protocol.

In an attempt to mitigate these problems, I do two things. First, an EU Directive requiring catalytic converters in all new vehicles was introduced after the Sofia meeting. A similar directive was shortly thereafter introduced in Iceland (in 1995). As this could be interpreted as policy spillovers, I exclude Iceland from the donor pool in the Sofia and the Geneva protocols. Second, I use several exclusion criteria on the donor pool to see how this influences estimated treatment effects. In the baseline estimation, I keep all non-ratifying LRTAP countries in the donor pool, i.e., countries that have ratified the initial 1979 Convention and potentially other LRTAP-protocols, but not the specific protocol in question.[50] If positive spillovers are substantial and complementarities are strong, we would expect treatment effects to be underestimated. In robustness checks, I make several changes to the donor pool, such as (i) removing all LRTAP countries from the donor pool, and (ii) restricting the donor pool to LRTAP-countries only.

5.2.2 Choice of treatment window

Previous studies have typically used each country's ratification year as the intervention date. This might be problematic if there are anticipation effects, or if certain components were in place prior to the formal implementation. In the case of the Helsinki protocol, the first LRTAP meeting in November 1979 represented the start of the international cooperation efforts. The

[49]The majority of SO_2 emissions and NO_x emissions stem from combustion of fossil fuels. By contrast, VOCs are emitted from a wide range of sources, including household and office products, loading, storage and transportation of crude oil, and road traffic. Road transport is also a major source of NO_x, linking the two pollutants. Further, when NO_x and VOCs are exposed to sunlight, they are transformed into ground-level ozone, which has adverse health effects. This implies that initiatives to combat ozone might have an effect on both pollutants.

[50]I make an exemption for LRTAP countries that have signed, but not ratified the protocol in question, which I exclude from the donor pool. The Geneva protocol is the only protocol where a few countries have signed but not ratified the protocol.

primary focus of the first meeting was SO_2 emissions, and we can therefore expect there to be anticipation effects for this pollutant. Further, the baseline year for the Helsinki protocol was 1980. If this was known in advance of the meeting in 1985, it could give an incentive for early reductions. To capture the full effect of the Helsinki protocol, I therefore define the treatment date to be the year after the first Convention meeting (1980).[51]

For NO_x, I define the intervention to the year before the Sofia meeting, i.e., 1987. This is the year used as the baseline for emission reductions in the protocol. We might, however, see an effect already from 1980 due to complementaries between SO_2 and NO_x. As there is a weaker link between VOCs and the two other pollutants, we might expect smaller effects of the first Convention and the two subsequent protocols on emissions of VOCs. Still, there might be an anticipation effect. In the same way as for NO_x, I define the intervention as the year before the Geneva meeting, i.e., 1990, which is (one of) the baseline year(s) used in the protocol text.

Lastly, I define the end year for each protocol as the year a new protocol was introduced to replace the old one. The Oslo protocol opened up for signatures in 1994, and was meant to replace the Helsinki protocol. The Gothenburg protocol replaced all previous protocols on SO_2, NO_x and VOCs, and opened up for signatures in 1999. This implies the following treatment windows for the three protocols: 1980-1994 for the Helsinki protocol, 1987-1999 for the Sofia protocol, and 1990-1999 for the Geneva protocol.[52]

5.2.3 Choice of outcome variable

A potential problem in assessing the effects of the protocols is the fact that LRTAP countries often have higher pollution levels than non-LRTAP countries, which might make it hard to find a good match in the donor pool. As a way around this problem, I normalize emissions per capita by setting emissions to be equal to 100 in the year before the intervention.[53] Normalizing emissions eases comparability of countries, and also lets me pool estimates to arrive at an average

[51] As data on GDP is only available from 1980 and onwards for many countries, it is more convenient to use 1980 than 1979 as the treatment date. Also, as the meeting found place in November 1979, 1980 might better reflect the timing of the intervention. To test the sensitivity of the results to the choice of treatment date, I also use the year the Helsinki protocol entered into force (1987) as an alternative intervention date (see Section 6.5).

[52] Alternative approaches could be to (i) use the end year of the emission targets (i.e., 1993 for the Helsinki protocol, 1994 for the Sofia protocol, and 1999 for the Geneva protocol), or (ii) use the year the Oslo and Gothenburg protocols entered into force (i.e., 1998 and 2005, respectively).

[53] A similar approach is taken in Cavallo et al. (2013) and Almer and Winkler (2017).

treatment effect. For the transformation to make sense, I need to assume that donor countries with lower pollution levels are able to reproduce trends in emissions for treated countries with a higher pollution level.

5.2.4 Choice of predictors

Lastly, as country characteristics are not available for all countries in the pre-treatment period, I face a trade-off in which variables to include as predictors. In the main specification, I include the following predictors: normalized emissions per capita in the years prior to the intervention, emissions per capita in the treatment year, GDP per capita in the treatment year, and the share of fossil fuel of total energy use in the treatment year.[54]

5.3 Final sample

As some of the restricting criteria presented above are pollutant-specific, the donor pool will be slightly different for the three pollutants. Further, as the treatment group is defined as countries signing and ratifying the protocol in question, the number of treated countries will also vary. Appendix Table C.3 lists countries used for estimating the effect of the Helsinki, Sofia and Geneva protocols. The donor pools used in the main analysis consists of 43-51 countries.

5.4 Descriptives

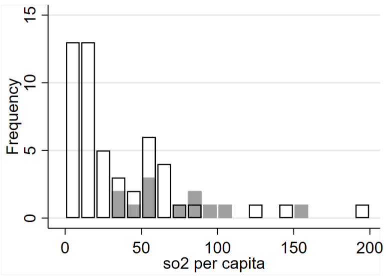
Figure 1 shows the distribution of emission levels per capita and country characteristics, by treated and donor countries. Although treated countries tend to have higher emission levels per capita, there is common support (i.e., an overlap) for all three pollutants. This implies that it should be possible to construct a synthetic control that closely matches the emission level of the treated country. The same is true for GDP growth, population growth and the share of fossil fuel of total energy use. For GDP per capita, however, there is limited common support for the richest countries in the treatment group, such as Switzerland and Norway. This means that it

[54]Appendix Table A.1 shows which LRTAP countries lack data on GDP. The relative importance of each predictor (v_m in equation 2) are set to the following values, based on information from several test runs: normalized emissions: 0.5, emission levels: 0.4, GDP: 0.095, and fossil share: 0.005. By fixing the weights v_m, I ensure that predictors are weighted in the same way for all countries.

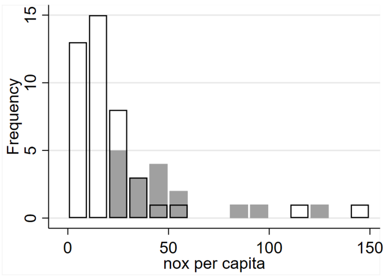
will not be possible to construct a synthetic control that exactly reproduces the income level of these countries.

Figure 1: Histograms. 1980

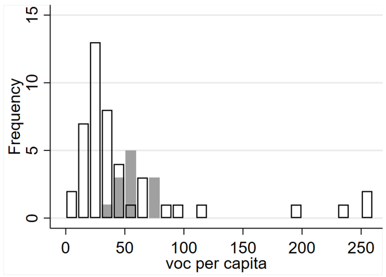
(a) SO_2 per capita



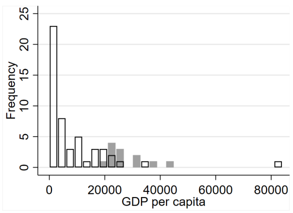
(b) NO_x per capita (1987)



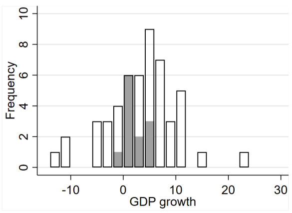
(c) VOC per capita (1990)



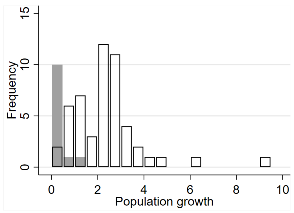
(d) GDP per capita



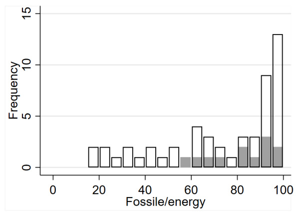
(e) GDP growth



(f) Population growth



(g) Fossil/energy



Notes: Figures show histograms of variables used as predictors. The vertical axes indicate the number of countries. Gray bars indicate treated countries and hollow bars indicate donor countries. The sample is the one used for estimating the effect of the Helsinki protocol - with the exception of panel (b) and (c), where the samples for the Sofia (b) and Geneva (c) protocols are used. Unless stated otherwise, statistics are based on data from the year 1980. See Appendix Table C.4 for means and standard deviations.

6 Results

In the following, I report results from the synthetic control method, where countries that have signed and ratified the protocol in question are defined as the treatment group.[55] The outcome variables are normalized values of SO_2, NO_x and VOC emissions per capita.

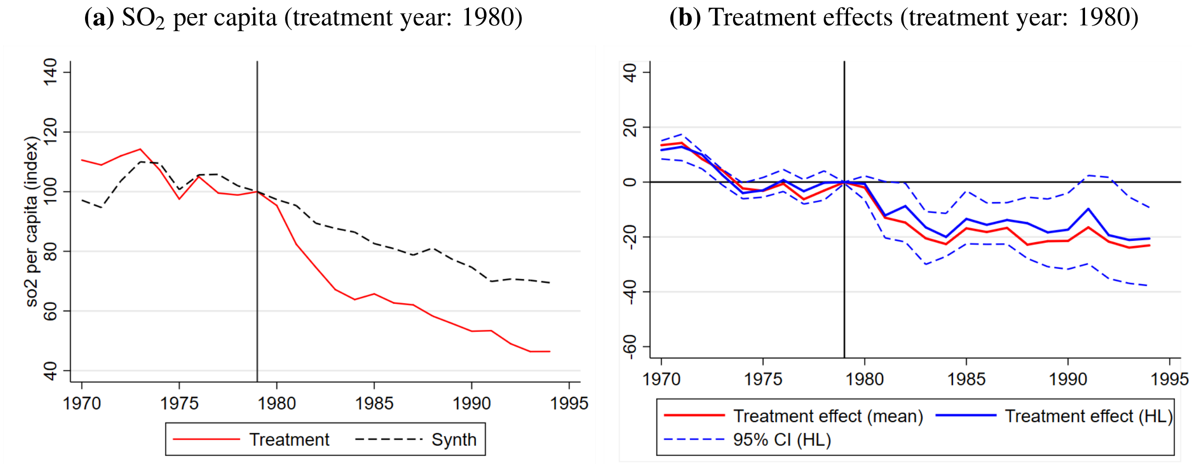
6.1 Effects of the Helsinki protocol on SO_2 emissions

Figure 2 and Table 3 summarize the effects of the Helsinki protocol on SO_2 emissions. Using the year after the LRTAP Convention as the intervention year (1980), countries that ratified the Helsinki protocol experienced a sharp decline in emissions in the post-treatment period, see Figure 2a. The rate of decline is larger for the treated countries than for the synthetic control.

[55] To estimate treatment effects, I use the *synth* package in STATA, developed by Abadie et al. (2010).

From Figure 2b we observe that treatment effects are significantly different from zero in the post-treatment period.[56] On average, SO_2 emissions reductions were 17-18% larger compared to the synthetic control in the treatment period (see Table 3) and the effect is significant at a 1% level when using a two-sided test. Looking at the development over time, most of the reductions materialize in the first 5 years. After 5 years, emissions are 20-23% lower than the synthetic control, and emissions fluctuate around this level for the rest of the period.

Figure 2: Effects of the Helsinki protocol on SO_2 emissions



Notes: Panel (a) shows the development in emissions for the treatment group (red line) and the synthetic control group (black, dashed line). Emissions in the year before treatment is normalized to 100. Panel (b) shows yearly treatment effects. The solid red line corresponds to the average, yearly treatment effects $\bar{\alpha}_{e1t}$ estimated from equation 5. The solid blue line indicates the HL point estimate (see Section 4.4 for details). The dashed blue lines indicate a 95% confidence interval. For weights used, see Appendix Table D.4.

Table 3: Pooled treatment effect for SO_2. 1980-1994.

	Average	Specific years	
	1980-1994	Year 5	Year 10
Treatment effect (mean)	-18.371	-22.603	-21.543
Treatment effect (HL)	-16.583	-20.000	-18.333
Mean rank	0.294 ***	0.236 ***	0.288 ***
95% CI (low)	-23.750	-27.083	-27.083
95% CI (high)	-6.417	-11.417	-11.417

Notes: Critical values for the mean percentile rank are derived from the simulation procedure described in Appendix B.2. Inverting this rank gives the 95% confidence intervals (CI). See Table D.1 for yearly treatment effects. Significance level: 1%: .302 , 5%: .349 , 10%: .375
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

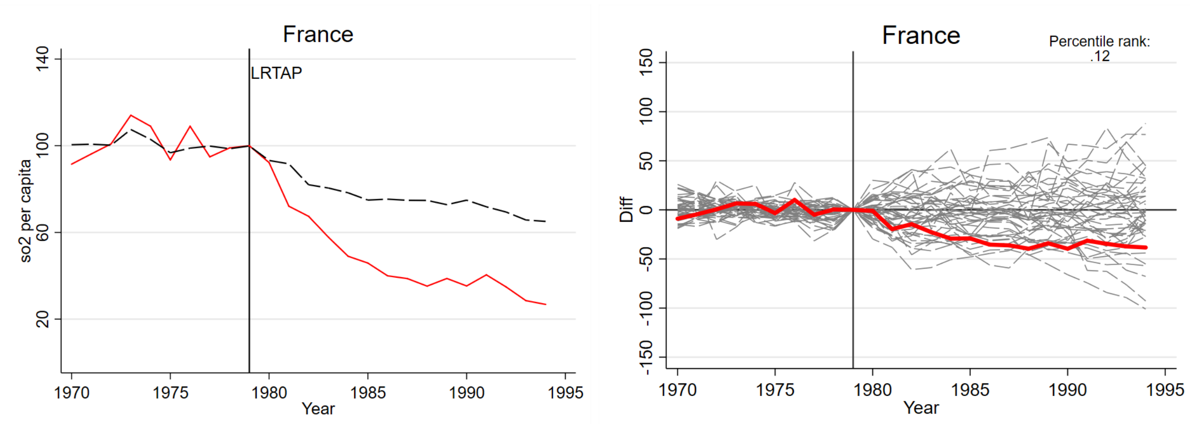
[56] The Hodges-Lehman (HL) point estimate corresponds to the mean of the upper and lower confidence bands, see Section 4.4.

6.1.1 Country estimates

The pooled treatment estimate is based on individual country estimates. Figure 3a illustrates the development in emissions for a single treated country (France) and the corresponding synthetic control. The figure shows a sharp decline in SO_2 emissions after 1980. This rate of decline is much larger than for the "synthetic France". To evaluate the statistical significance of this difference, I run placebo treatments on the countries in the donor pool. These are presented in Figure 3b. Ranking treatment effects from lowest to highest, France has the fifth largest decline in emissions relative to the synthetic control. With 42 countries in the donor pool, this results in a percentile rank of around 12%.[57]

Figure 3: Development in SO_2 emissions for France and the synthetic control.

(a) Treatment and synthetic control **(b)** Treatment effects and placebo runs

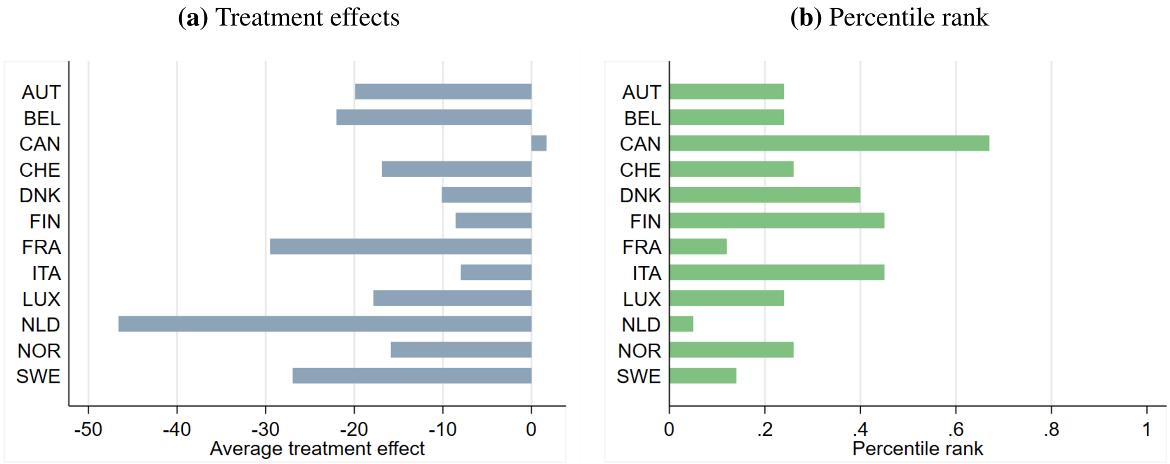


Notes: Panel (a) shows the development in emissions for the treated country (red line) and the synthetic control (black, dashed line). Emissions in the year before treatment is normalized to 100. Panel (b) shows the difference between the treated and synthetic control outcomes $\hat{\alpha}_{1t}$ from equation 3 (thick, red line) and placebo runs (thin, gray lines). After trimming the donor countries by the MSPE, there are 42 countries left in the donor pool. See the Online Appendix for similar figures of all treated countries.

Looking at all countries ratifying the Helsinki protocol, only one country (Canada) experienced a non-negative treatment effect, see Figure 4a. While the individual treatment effects are never statistically significant at a 5% level when using a two-sided test (see Figure 4b), the pooled treatment effects is significantly different from zero at a 1% level (see Table 3).

[57]Using a one-sided test, the percentile rank corresponds to a significance level of 12%. Using a two-sided test, the percentile rank corresponds to a significance level of 24%.

Figure 4: Country-specific treatment effects. SO_2. 1980-1994



Notes: Panel (a) shows the average, country-specific treatment effect derived from equation 4. Panel (b) shows the country-specific, percentile ranks derived from equation 6. See Table F.1 in the Online Appendix for country-specific treatment effects in table format. See Figure F.1 the Online Appendix for figures showing the country-specific treatment effects over time and placebo runs.

6.1.2 Match quality and weights

The validity of the estimated treatment effects depends on how closely the synthetic controls approximates key predictors for the treated countries. From Figure 2, we see that the pre-intervention development in normalized emissions is not a perfect match. However, from Figure 2a and Figure 2b we see that the 5-6 years before the treatment show a good match.

In addition to normalized emissions, I use pollution per capita, GDP per capita and fossil fuel share as predictors. Table 4 shows the average match of the four predictors.[58] While the yearly development in normalized emissions show some deviations, the average over the pre-treatment period is very similar. For SO_2 emissions per capita and GDP per capita, the average values are somewhat higher for the treated countries.[59] However, the difference in pollution levels is almost entirely driven by one country: Canada. For all other countries, the pollution level per capita is about the same (see Appendix Figure D.1). For GDP per capita, the difference is driven by five countries: Canada, Switzerland, Denmark, Luxembourg, and Norway. If we restrict the treatment group to countries with a close match on GDP, the average treatment effect actually increases to -23%. The favorable treatment effect is hence not driven by the richest countries in the sample.

[58] Figure D.1 in the Appendix shows the match of predictors for each individual country.

[59] Note that the difference is much smaller compared to the average of the entire donor pool (see Table C.4)

An overview of the weights assigned to each country in the donor pool is provided in Appendix Table D.4. The countries most often used to construct synthetic controls are the United States, Iceland, New Zealand, and the United Kingdom.

Table 4: Match of predictors. SO_2. 1980.

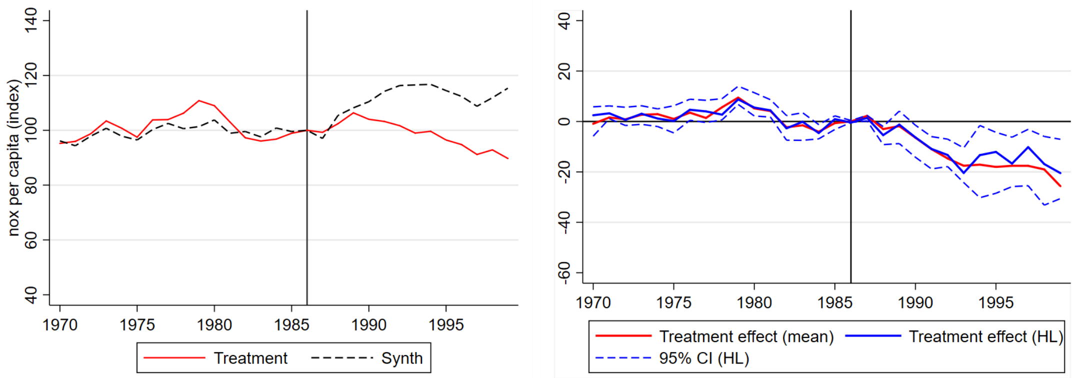
Predictor	Treated	Synthetic
SO2 per capita (normalized) (1970-1980)	104.42	102.42
SO2 per capita (level) (1980)	73.92	69.83
GDP per capita (1980)	27787.75	24320.25
Fossil share (1980)	81.92	76.58

6.2 Effects of the Sofia protocol on NO_x emissions

The estimated effect of the Sofia protocol on NO_x emissions is reported in Figure 5 and Table 5. Using the year before the Sofia meeting as the treatment year (1987), the synthetic control method shows significant treatment effects in the range of 12-13%. In other words: NO_x emissions in the treated countries were on average 12-13% lower compared to the synthetic control. The treatment effect is significant at a 1% level. Five years after the intervention, emissions were around 11% lower than the synthetic control, while this difference increases to 17-18% after ten years.

Figure 5: Effects of the Sofia protocol on NO_x emissions

(a) NO_x per capita (treatment year: 1987) **(b)** Treatment effects (treatment year: 1987)



Notes: Panel (a) shows the development in emissions for the treatment group (red line) and the synthetic control group (black, dashed line). Emissions in the year before treatment is normalized to 100. Panel (b) shows yearly treatment effects. The solid red line corresponds to the average, yearly treatment effects $\bar{\alpha}_{e1t}$ estimated from equation 5. The solid blue line indicates the HL point estimate (see Section 4.4 for details). The dashed blue lines indicate a 95% confidence interval. For weights used, see Appendix Table D.5.

Table 5: Pooled treatment effect for NO_x.

	Average	Specific years	
	1987-1999	Year 5	Year 10
Treatment effect (mean)	-12.858	-10.966	-17.570
Treatment effect (HL)	-11.824	-10.882	-16.706
Mean rank	0.319 ***	0.301 ***	0.314 ***
95% CI (low)	-20.000	-18.824	-18.824
95% CI (high)	-5.765	-6.000	-6.000

Notes: Critical values for the mean percentile rank are derived from the simulation procedure described in Appendix B.2. Inverting this rank gives the 95% CIs. Average treatment effects by year are reported in Table D.2.
Significance level: 1%: .333 , 5%: .375 , 10%: .396
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The year the Sofia protocol opened up for signature (1988), a European Communities[60] Directive was introduced, which required large combustion plants to significantly reduce SO_2 and NO_x emissions compared to 1980 levels.[61] Furthermore, in 1993 a EU directive targeting cars was launched, requiring all cars sold within the European Union to be fitted with a catalytic converter, which lowers NO_x emissions.[62] Both Directives are in line with the goals stated in the Sofia Protocol, and could be interpreted as a result of the international cooperation. Specifically, Article 2 of the Sofia protocol states that countries need to introduce emissions standards or other pollution control measures to stationary sources, and Article 4 mandates countries to facilitate the circulation of vehicles equipped with catalytic converters.[63]

6.2.1 Country estimates

Figure 6 gives an overview of the country-specific treatment effects and corresponding percentile ranks. With the exception of four countries (Canada, Spain, Greece, and Luxembourg), the treatment group experienced a decrease in emissions compared to the synthetic control. Al-

[60]European Union (EU) from 1993.

[61]*The Council Directive 88/609/EEC of 24 November 1988 on the limitation of emissions of certain pollutants into the air from large combustion plants*. For more information, see `http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31988L0609`.

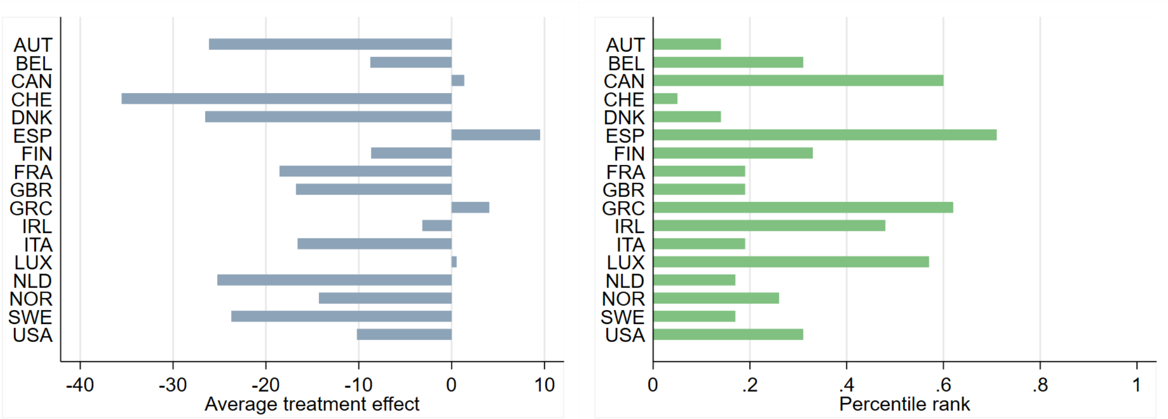
[62]The policy also diffused to other non-EU countries, and in 1995 Iceland required all new vehicles to have a catalytic converter.

[63]It could also be the case that the Sofia protocol simply reflected coordination efforts within the European Union that would have emerged also in the absence of the protocol. However, as the Directives were introduced after the Sofia protocol meeting, it might be reasonable to assume that the LRTAP framework contributed to the decision to implement the Directives.

though country estimates are never statistically significant at a 5% level (using a two-sided test), the pooled treatment effect is significantly different from zero at a 1% level (see Table 5).

Figure 6: Country-specific treatment effects. NO_x. 1987-1999

(a) Treatment effect **(b)** Percentile rank



Notes: Panel (a) shows the average, country-specific treatment effect derived from equation 4. Panel (b) shows the country-specific, rank-based p-values derived from equation 6. See Table F.4 in the Online Appendix for country-specific treatment effects in table format. See Figure F.0 the Online Appendix for figures showing the country-specific treatment effects over time and placebo runs.

6.2.2 Match quality and weights

The average development in NO_x emissions for the synthetic control tracks the treatment group fairly well in the pre-intervention period, see Figure 5b. Although the pre-treatment difference is significantly different from zero in some years, the treatment effect fluctuates around zero for the 16 year period prior to the intervention. Table 4 shows the average match of the four predictors.[64] Both normalized NO_x emissions and NO_x levels show a close match. GDP per capita, however, is higher for the treated countries, while the opposite is the case for the fossil fuel share. Countries frequently used as controls include Japan, United Arab Emirates, and Hong Kong.[65]

[64]Figure D.1 in the Appendix shows the match of predictors for individual countries.

[65]Note that more countries are included in the treatment group under the Sofia protocol compared to the Helsinki protocol, including the UK and the U.S. This means that these countries are no longer in the donor pool, and we are left with "less similar" countries. For information on which countries in the donor pool are used to construct synthetic controls, see Table D.5 in the Appendix.

Table 6: Match of predictors. NO_x. 1987.

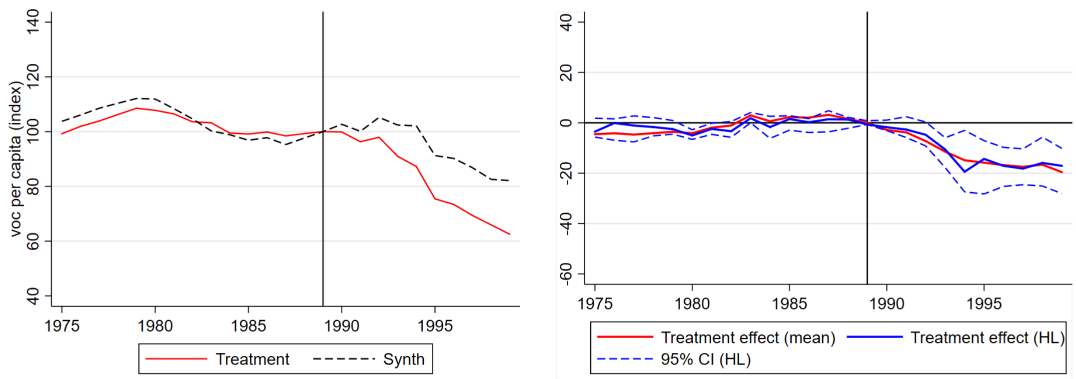
Predictor	Treated	Synthetic
NOX per capita (normalized) (1970-1987)	100.82	99.18
NOX per capita (level) (1987)	43.35	43.06
GDP per capita (1987)	28929.82	24395.41
Fossil share (1987)	77.47	86.53

6.3 Effects of the Geneva protocol on VOCs

Figure 7 depicts the estimated yearly effects of the Geneva protocol on emission of VOCs. Using the year before the meeting as the treatment date (1990), the development in emissions for the treatment group and the synthetic control diverge right after the intervention. From Figure 7b, we see that the average treatment effect increases over time. In the post-intervention period, emissions reductions for treated countries were on average 13-14 % larger than for the synthetic counterfactual (see Table 7), and the effect is significant at a 1% level. Five years after the intervention, emissions were 15-19% lower than the synthetic control. This difference is slightly larger ten years after the intervention (17-20%).

Figure 7: Effects of the Geneva protocol on emissions of VOCs

(a) VOCs per capita (treatment year: 1990) **(b)** Treatment effects (treatment year: 1990)



Notes: Panel (a) shows the development in emissions for the treatment group (red line) and the synthetic control group (black, dashed line). Emissions in the year before treatment is normalized to 100. Panel (b) shows yearly treatment effects. The solid red line corresponds to the average, yearly treatment effects $\bar{\alpha}_{e1t}$ estimated from equation 5. The solid blue line indicates the HL point estimate (see Section 4.4 for details). The dashed blue lines indicate a 95% confidence interval. For weights used, see Appendix Table D.5.

Table 7: Pooled treatment effect for VOC.

	Average	Specific years	
	1990-1999	Year 5	Year 10
Treatment effect (mean)	-12.644	-14.866	-19.599
Treatment effect (HL)	-14.000	-19.417	-17.083
Mean rank	0.254 ***	0.266 ***	0.262 ***
95% CI (low)	-20.167	-27.417	-27.417
95% CI (high)	-6.000	-3.000	-3.000

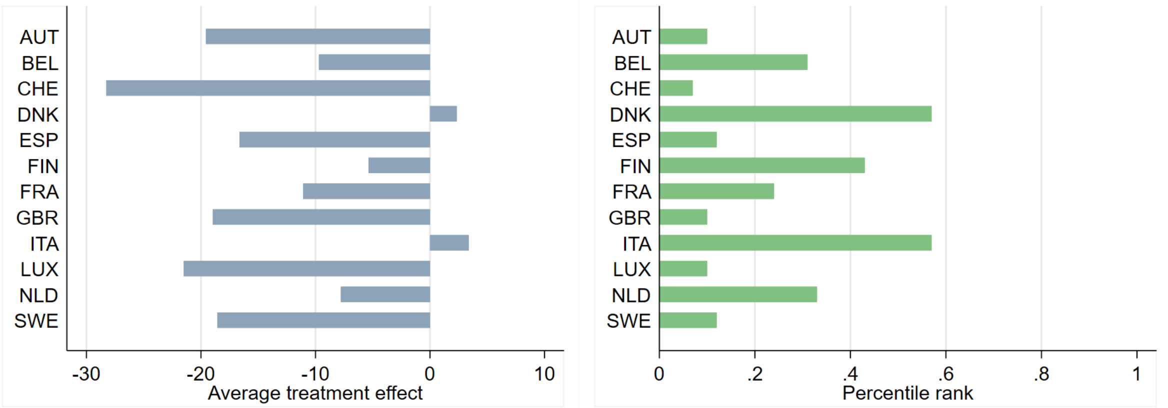
Notes: Critical values for the mean percentile rank are derived from the simulation procedure described in Appendix B.2. Inverting this rank gives the 95% CIs. Average treatment effects by year are reported in Table D.3.
Significance level: 1%: .302 , 5%: .349 , 10%: .375
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.3.1 Country estimates

Almost all countries experienced a decline in emissions relative to the synthetic control. Exceptions are Denmark and Italy. The percentile ranks are generally small - although none of the country-specific treatment effects are statistical significant at a 5 % level using a two-sided test.

Figure 8: Country-specific treatment effects. VOC. 1990-1999

(a) Treatment effect **(b)** Percentile rank



Notes: Panel (a) shows the average, country-specific treatment effect derived from equation 4. Panel (b) shows the country-specific, rank-based p-values derived from equation 6. See Table F.7 in the Online Appendix for country-specific treatment effects in table format. See Figure F.-2 the Online Appendix for figures showing the country-specific treatment effects over time and placebo runs. Norway is excluded from the treatment group due to an extreme development in VOCs.

6.3.2 Match quality and weights

Looking at Figure 7, we see that the synthetic control tracks the treatment group closely in the period prior to intervention. The average pollution level per capita in 1990 is also very similar for the treated countries and the synthetic controls, see Table 8. Again, it is hard to find a close match on the GDP predictor as ratifying countries are systematically richer than those in the donor pool. Japan is weighted heavily in the construction of the synthetic control countries.[66]

Table 8: Match of predictors. VOC. 1990.

Predictor	Treated	Synthetic
VOC per capita (normalized) (1975-1990)	102.25	103.42
VOC per capita (level) (1990)	51.17	49.75
GDP per capita (1990)	32521.33	26550.42
Fossil share (1990)	74.92	87.17

6.4 Comparing emission reductions to stated targets

How do emission reductions compare to the actual targets in each of the protocols? The Helsinki protocol committed ratifiers to reduce SO_2 emissions by at least 30% compared to 1980 levels, as soon as possible or by 1993. Actual emissions were 49% lower in 1993, while the constructed counterfactual show a reduction of 27%. This implies two things: first, that ratifying countries reduced emissions well beyond the target, and second, that emissions would have declined also in the absence of the Helsinki protocol, but not nearly as much. Out of the 49% decline in SO_2 emissions, around 22 percentage points was induced by the Helsinki protocol. For the Sofia protocol, the stated goal was to reduce NO_x emissions to 1987 levels by 1994. On average, actual emissions were at the same level in 1994 as in 1987, meaning that they were on target, while the synthetic counterfactual indicates a 20% increase in emissions in 1994 compared to 1987. The empirical findings hence suggest that the Sofia protocol staggered a rise in emissions. Under the Geneva protocol, countries could choose between different emission reduction targets, where one of them was a 30% reduction by 1999 compared to 1990. Actual emissions were 37.5% lower in 1999, while the synthetic control shows a 20.5% reduction, suggesting that the Geneva protocol contributed to an additional reduction of around 17 percentage points.

[66]For information on which countries are used to construct synthetic controls, see Appendix Table D.6.

6.5 Robustness checks

This section reports results from different robustness checks. Section 6.5.1 evaluates the statistical significance of the main results under alternative inference procedures. Section 6.5.2 makes changes to the predictor set or donor pool.[67] Section 6.5.3 uses later intervention dates. Section 6.5.4 reports results from a DiD approach to evaluate impacts of the Helsinki protocol.

6.5.1 Using alternative inference procedures

Table 9: Comparing significance levels under different inference procedures

	Helsinki (1980-1994)	Sofia (1987-1999)	Geneva (1990-1999)
Treatment effect (mean)	-18.371	-12.858	-12.644
Treatment effect (HL)	-16.583	-11.824	-14.000
Mean rank	0.294	0.319	0.254
Significance levels under different procedures:			
Independent ranks (Irwin-Hall, discrete)	1%	1%	1%
Independent ranks (Irwin-Hall, continuous)	5%	1%	1%
Independent ranks (randomization)	1%	1%	1%
Dependent ranks (randomization)	10%	1%	1%

Notes: Significance levels are from a two-sided test. See Sections B.1–B.4 in the Appendix for details on the inference procedures. All cut-off values are listed in Table B.3

Table 9 shows the statistical significance of the main results under different inference procedures.[68] The first procedure is the one used in the main analysis (*Independent ranks (Irwin-Hall, discrete)*). The second procedure uses continuous values instead of discrete to construct cut-off values, and results in a significance level of 5% instead of 1% in the case of the Helsinki protocol. The third approach randomly permutes the treatment status in a dataset with both actually treated countries and donor countries (*Independent ranks (randomization)*). Using this alternative inference procedure results in the same significance levels as in the main analysis for all three protocols. The fourth approach accounts for rank dependency by constraining the "treatment group" to consist of countries located geographically close to each other (*Dependent ranks (randomization)*). For the Helsinki protocol, the procedure generates a significance level

[67] See Section G in the Online Appendix for country weights chosen in the robustness checks.

[68] The inference procedures are described in Section 4.2.2, as well as in Sections B.3 and B.4 in the Appendix.

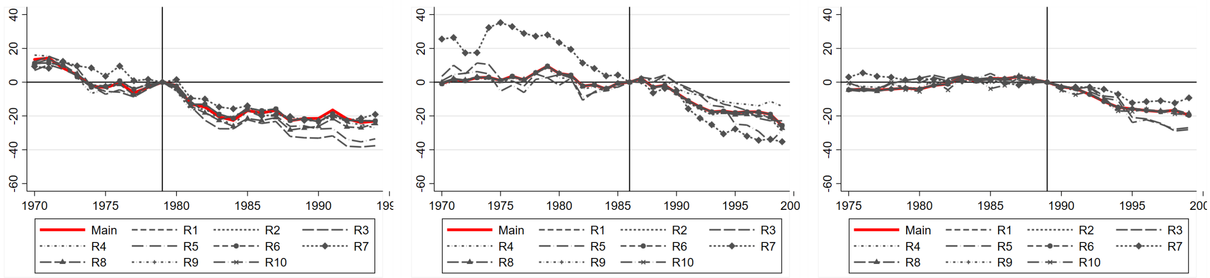
of 10%, while the estimated effect of the other two protocols are still significant at a 1% level.

6.5.2 Changing the predictor set and donor pool

Figure 9 and Table 10 summarize results from ten different robustness checks. In robustness checks R1-R5, I make changes to the predictor set, while in robustness checks R6-R10, I make changes to the donor pool. Overall, all robustness checks give negative treatment effects and produce relatively similar results as the main specification.

Figure 9: Robustness checks

(a) Helsinki (SO_2) **(b)** Sofia (NO_x) **(c)** Geneva (VOCs)



Notes: Each figure shows average, yearly treatment effects $\bar{\alpha}_{e1t}$ estimated from equation 5 for ten different robustness checks. For a description of the robustness checks see Table 10 and the text. See Appendix G for weights used in each robustness check.

Looking at the Helsinki protocol, treatment effects are either similar or slightly larger when changing the predictor set (see R1-R5). Excluding GDP as a predictor has the largest impact on the estimated treatment effect, increasing it by several percentage points (see R3 and R5 in Table 10). This is to some extent expected, as lower income countries at a different stage of economic development might now be used as controls. For the Sofia protocol, making changes to the predictor set either has no effect or slightly reduces the treatment estimate. The latter is the case in columns R3 and R4, where the fossil fuel share and the GDP level or the pollution level is jointly excluded. The estimated effect of the Geneva protocol is relatively stable across the different specifications.

Next, I make changes to the donor pool. I start by excluding all non-ratifying LRTAP countries from the set of potential control countries. As a result, the (absolute) treatment effect of the Helsinki protocol increases from -18% to -28% (see R6), while the HL estimate is less effected. The larger treatment effect could potentially be due to favorable spillovers across LRTAP countries, which would underestimate the true treatment effect if these countries are included in the

Table 10: Robustness checks

		Changing predictors					Changing donor pool				
	Baseline	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
Panel A: Helsinki (SO_2)											
Treatment effect (mean)	-18.371	-18.106	-18.917	-27.362	-19.850	-22.667	-27.799	-16.726	-22.061	-18.230	-19.979
Treatment effect (HL)	-16.583	-11.083	-15.417	-24.833	-14.167	-21.833	-19.167	.	-18.000	-13.583	-18.500
Mean rank	0.294 ***	0.317 **	0.280 ***	0.173 ***	0.294 ***	0.258 ***	0.252 ***	.	0.229 ***	0.335 **	0.282 ***
Panel B: Sofia (NO_2)											
Treatment effect (mean)	-12.858	-14.176	-13.975	-9.426	-7.619	-13.918	-13.144	-20.791	-13.696	-13.149	-15.371
Treatment effect (HL)	-11.824	-10.647	-14.059	-8.882	-2.471	-11.118	-12.471	.	-11.471	-13.471	-13.882
Mean rank	0.319 ***	0.338 **	0.280 ***	0.381 *	0.464	0.335 **	0.307 ***	.	0.337 **	0.249 ***	0.300 ***
Panel C: Geneva (VOCs)											
Treatment effect (mean)	-12.644	-12.634	-12.676	-14.983	-13.245	-15.614	-12.444	-7.675	-12.291	-12.349	-8.804
Treatment effect (HL)	-14.000	-15.667	-12.833	-16.000	-13.083	-11.500	-18.000	.	-13.583	-13.667	-11.083
Mean rank	0.254 ***	0.224 ***	0.242 ***	0.206 ***	0.232 ***	0.246 ***	0.222 ***	.	0.225 ***	0.286 ***	0.349 **

Notes: Critical values for the mean percentile rank are derived from the simulation procedure described in Appendix B.2. Inverting this rank gives the 95% CIs. Treatment year SO_2: 1980. Treatment year NO_x: 1987. Treatment year VOC: 1990. For cut-off values, see Appendix Table B.3. For weights, see Appendix Section G.
R1: Baseline sample. Add GDP growth and population growth to predictor set.
R2: Baseline sample. Drop fossil fuel share from predictor set
R3: Baseline sample. Drop fossil fuel share and GDP per capita from predictor set
R4: Baseline sample. Drop fossil fuel share and pollution per capita from predictor set
R5: Baseline sample. Only keep normalized pollution and fossil fuel share in the predictor set
R6: Baseline predictor set. Drop non-ratifying LRTAP countries from donor pool.
R7: Baseline predictor set. Only keep non-ratifying LRTAP countries in the donor pool.
R8: Baseline predictor set. Drop countries in the bottom 40% of the income distribution
R9: Baseline predictor set. Keep countries of all income levels, as well as countries experiencing highly volatile emissions
R10: Baseline predictor set. Drop (literately) the top two donor countries (SO2: United States and New Zealand, NOx: Japan and New Zealand, VOC: Japan and Ireland).
Significance levels for SO_2: 1%: .302 , 5%: .349 , 10%: .375
Significance levels for NO_x: 1%: .333 , 5%: .375 , 10%: .396
Significance levels for VOCs: 1%: .302 , 5%: .349 , 10%: .375
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

donor pool. Countries often assigned a large weight in the construction of synthetic controls include Australia, New Zealand, Japan, and Chile (see Appendix G). Treatment effects of the Sofa and Geneva protocols are not notably affected, which might not be surprising as non-ratifying LRTAP countries make up a small part of the donor sample.

In R7, I do the opposite of R6 and restrict the donor pool to non-ratifying LRTAP countries only. This lowers the treatment effect of the Helsinki protocol slightly, from -18.3% to -16.7%. The same is the case for the Geneva protocol, where the estimated treatment effect is around 5 percentage points lower. Given the potential presence of favorable spillover effects, lower treatment effects are to be expected. The estimated effects of the Sofia and Geneva protocols rely almost entirely on 1-2 donor countries (Cyprus and Ireland), which means that any post-intervention shocks or spillovers to these countries would severely impact the treatment estimates. The pre-treatment match for the Sofia protocol is also poor, meaning that the estimated (larger) treatment effect is not reliable. Due to the small number of countries in the donor pool, I cannot perform inference on these estimates.

As shown in Section 5.2, GDP per capita is systematically higher for treated countries compared to donor countries. In the main estimation, I try to mitigate this problem by including GDP per capita as a predictor, and also drop the poorest quintile from the donor pool. In column R8, I adjust the donor pool by excluding the *two* poorest quintiles. This increases the effect of the Helsinki protocol somewhat (from -18.3% to -22%), while the effects of the Sofia and Geneva protocol are very similar to the baseline. The small changes to the estimated effects are not surprising as low-income countries are rarely given a positive weight in the baseline estimation. In R9, I instead expand the donor pool by including countries of all income levels as well as countries experiencing highly volatile emissions. This has little effect on the treatment estimates, but lowers the statistical significance level of the Helsinki protocol from 1% to 5%.

Overall, the countries used to construct the synthetic controls vary substantially across the different robustness checks.[69] This is reassuring, as it suggests that the constructed counterfactuals are not just an artifact of a specific country combination. However, we might still worry that the synthetic controls rely on a couple of key donor countries, which could make the treat-

[69] See Appendix G for weights used in each robustness check.

ment estimates vulnerable to specific shocks to these countries[70] In R10, I exclude (iteratively) the top two donor countries used to construct synthetic controls in each of the three protocols.[71] This slightly increases estimated effects of the Helsinki and Sofia protocols, and slightly lowers the effect of the Geneva protocol.

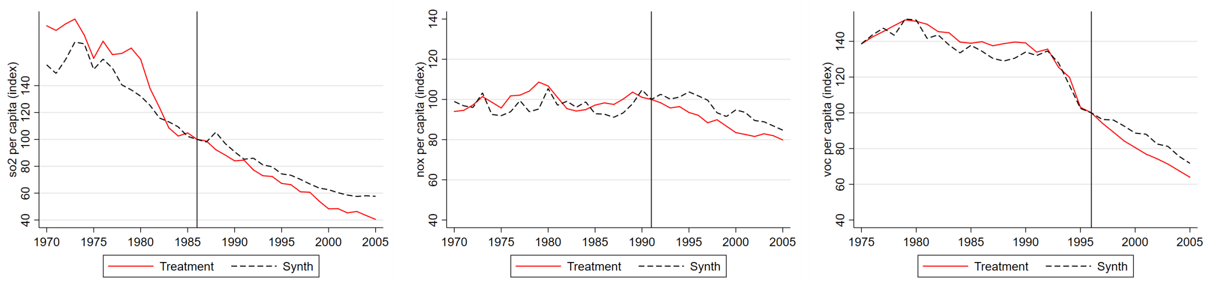
To sum up, estimated treatment effects are fairly robust to the 10 different robustness checks, and all estimates show a negative effect on emissions. I also show that drastically reducing the donor pool to non-ratifying LRTAP countries only still renders negative treatment effects, but of a lower magnitude than the baseline results.

6.5.3 Changing the intervention date

Figure 10 shows results for the three protocols when delaying the intervention date. For Helsinki, I use the year the protocol entered into force (1987) as an alternative intervention date. When delaying the intervention, there is no longer a significant effect of the protocol, see Figure 10a. For the Sofia protocol, I change the intervention year from 1987 to 1992 and for the Geneva protocol, I change the intervention year from 1990 to 1997. The new dates reflect the time when the protocols entered into force. Delaying the interventions result in lower treatment effects, see Figures 10b and 10c. These results highlight the importance of defining an intervention year that accounts for anticipation effects in order to capture the full extent of the treatment.

Figure 10: Effects of the three protocols when delaying intervention

(a) Helsinki (treatment date: 1987) **(b)** Sofia (treatment year: 1992) **(c)** Geneva (treatment year: 1997)



Notes: Panels show the development in emissions for the treatment group (red line) and the synthetic control group (black, dashed line). Emissions in the year before treatment is normalized to 100.

[70] See Section D.2 in the Appendix.

[71] SO_2: The Unites States and New Zealand, NO_x: Japan and New Zealand, VOCs: Japan and Ireland.

6.5.4 Using a difference-in-differences approach

In Appendix E, I estimate effects of the Helsinki protocol on log SO_2 using a difference-in differences (DiD) approach. Results from a DiD model with leads and lags show that the assumption of parallel trends does not hold. Using 1985 as the treatment year instead of 1980 aggravate the discrepancy in pre-treatment trends, and also lowers the estimated treatment effect. Following Aakvik and Tjøtta (2011) and including country-specific linear time trends wipes out the treatment effect completely. The results imply that we clearly need to address the parallel trends assumption, but including linear time trends may not be a suitable approach as it seems to absorb most of the treatment effect.

7 Concluding remarks

Understanding the potential of international agreements to mitigate cross-border environmental externalities is crucial to guide policy makers towards instruments that actually make a difference. At the same time, evaluating impacts of multilateral agreements is methodologically challenging, and to date there are few empirical studies that credibly establish causal relationships between ratification status and subsequent environmental outcomes, such as air pollution.

In this study, I revisit three large-scale pollution protocols on SO_2, NO_x and VOCs implemented in the 1980s and early 1990s, with the aim is to establishing causal impacts of the protocols on emissions. By combining a newly assembled dataset with a method for constructing synthetic counterfactuals, I am able to address several potential problems associated with previous empirical examinations, such as non-parallel emission trends, spillovers, and anticipation effects. Results from the empirical estimation suggest that the international protocols induced sizable emissions reductions of all three pollutants. For ratifying countries, SO_2 emissions were on average 22% lower than the synthetic control group ten years after the intervention, while the corresponding numbers for NO_x and VOC emissions were 18% and 20%, respectively.

My findings suggest that international protocols can be an effective tool to induce countries to lower their emissions. This contrasts the often gloomy predictions from the game theoretical literature. The findings also illustrate how different pre-intervention trends and anticipation effects can bias estimated treatment effects, if these are not accounted for in a suitable manner. Specifically, the comparative analysis in this paper suggests that previous studies on the Helsinki protocol have tended to underestimate the favorable effects on SO_2 emissions.

Although I address several shortcomings of previous studies, there are still other, more inherent features of international protocols that make it challenging to recover causal estimates. In particular, accounting for all types of direct and indirect spillovers in a highly complex and globalized economy, is close to impossible. If ratifying countries lower their emissions by e.g., reducing the use of high sulfur coal, this could potentially lead to increased emissions in non-ratifying countries via trade flows. At the same time, higher technology adoption by ratifying countries might stimulate technological development and diffusion, potentially inducing emissions reductions also in non-ratifying countries. In order to fully account for all positive and negative spillovers, we would need detailed, global data on technology adoption and product-level trade flows between countries. Such an analysis is beyond the scope of this paper, but might be a fruitful avenue for future research.

References

Aakvik, A. and Tjøtta, S. (2011). Do collective actions clear common air? The effect of international environmental protocols on sulphur emissions. *European Journal of Political Economy*, 27(2):343–351.

Abadie, A. (2012). Using synthetic controls to evaluate an international strategic positioning program in uruguay: Feasibility, data requirements, and methodological aspects. Harvard University, mimeo.

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505.

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

Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque country. *American Economic Review*, 93(1):113–132.

Almer, C. and Winkler, R. (2017). Analyzing the effectiveness of international environmental policies: The case of the Kyoto protocol. *Journal of Environmental Economics and Management*, 82:125 – 151.

Barrett, S. (1994). Self-enforcing international environmental agreements. *Oxford Economic Papers*, pages 878–894.

Benchekroun, H. and Long, N. V. (2012). Collaborative environmental management: A review of the literature. *International Game Theory Review*, 14(04):1240002.

Billmeier, A. and Nannicini, T. (2013). Assessing economic liberalization episodes: A synthetic control approach. *Review of Economics and Statistics*, 95(3):983–1001.

Bratberg, E., Tjøtta, S., and Øines, T. (2005). Do voluntary international environmental agreements work? *Journal of Environmental Economics and Management*, 50(3):583–597.

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

Dinda, S. (2004). Environmental kuznets curve hypothesis: A survey. *Ecological Economics*, 49(4):431 – 455.

Dube, A. and Zipperer, B. (2015). Pooling multiple case studies using synthetic controls: An application to minimum wage policies. *IZA DP No. 8944*.

EPA (2015a). Reducing acid rain. Retrieved from: www.epa.gov/airquality/peg_caa/acidrain.html. Accessed: August 2015.

EPA (2015b). Sulfur dioxide. Retrieved from: http://www.epa.gov/airquality/sulfurdioxide/. Accessed: August 2015.

Finus, M. and Tjøtta, S. (2003). The Oslo protocol on sulfur reduction: the great leap forward? *Journal of Public Economics*, 87(9-10):2031–2048.

Gobillon, L. and Magnac, T. (2016). Regional policy evaluation: Interactive fixed effects and synthetic controls. *Review of Economics and Statistics*, 98(3):535–551.

Hodges Jr, J. L. and Lehmann, E. L. (1963). Estimates of location based on rank tests. *The Annals of Mathematical Statistics*, pages 598–611.

Hoel, M. (1992). International environment conventions: the case of uniform reductions of emissions. *Environmental and Resource Economics*, 2(2):141–159.

Houghton, K. A. and Naughton, H. T. (2016). *Comparative Law and Economics*, chapter Chapter 18: International environmental agreement effectiveness: A review of empirical studies. Elgar.

Jørgensen, S., Martín-Herrán, G., and Zaccour, G. (2010). Dynamic games in the economics and management of pollution. *Environmental Modeling & Assessment*, 15(6):433–467.

JRC (2012). The emissions database for global atmospheric research (EDGAR). European Commission, Joint Research Centre (JRC)/PBL Netherlands Environmental Assessment Agency. Available at: edgar.jrc.ec.europa.eu/datasets_list.php?v=42. Version 4.2.

Mideksa, T. K. (2013). The economic impact of natural resources. *Journal of Environmental Economics and Management*, 65(2):277 – 289.

Mitchell, R. B. (2002-2015). International environmental agreements database project (version 2014.3). Available at: http://iea.uoregon.edu/. Date accessed: 20 April 2015.

Murdoch, J. C., Sandler, T., and Sargent, K. (1997). A tale of two collectives: sulphur versus nitrogen oxides emission reduction in Europe. *Economica*, 64(254):281–301.

Murdoch, J. C., Sandler, T., and Vijverberg, W. P. M. (2003). The participation decision versus the level of participation in an environmental treaty: a spatial probit analysis. *Journal of Public Economics*, 87(2):337–362.

Naughton, H. T. (2010). Globalization and emissions in Europe. *The European Journal of Comparative Economics*, Vol. 7(No. 2):pp. 503–519.

OECD (2014). *OECD Factbook 2014: Economic, Environmental and Social Statistics*, chapter Sulphur and nitrogen emissions. OECD Publishing, Paris.

O'Neill, S., Kreif, N., Grieve, R., Sutton, M., and Sekhon, J. S. (2016). Estimating causal effects: considering three alternatives to difference-in-differences estimation. *Health Services and Outcomes Research Methodology*, 16(1-2):1–21.

Pinotti, P. (2015). The economic costs of organised crime: Evidence from southern italy. *The Economic Journal*, 125(586):203–232.

Powell, D. (2017). Synthetic control estimation beyond case studies: Does the minimum wage reduce employment?

Ringquist, E. J. and Kostadinova, T. (2005). Assessing the effectiveness of international environmental agreements: The case of the 1985 Helsinki protocol. *American Journal of Political Science*, 49(1):86–102.

Stern, D. I. (2006). Reversal of the trend in global anthropogenic sulfur emissions. *Global Environmental Change*, 16(2):207 – 220.

The World Bank (2015). The World Bank Indicators. Available at: http://data.worldbank.org/indicator/. Data accessed: July 2015.

UNECE (2015). Protocols. Retrieved from: www.unece.org/env/lrtap/status/lrtap_s.html. Accessed: August 2015.

Online appendices

Contents

Appendix A Additional information on the LRTAP

A.1 Ratification of international pollution protocols

Table A.1: Ratification of international protocols. Complete list

Country name	**LRTAP**	**Helsinki (SO_2)**	**Sofia (NO_X)**	**Geneva (VOC)**	**Oslo (SO_2)**	**Gothenburg (SO_2,NO_X, VOC)**	**GDP**	**Structural**
Armenia	1997						No	Yes
Austria	1982	1987	1990	1994	1998			
Belarus	1980	1986	1989				No	Yes
Belgium	1982	1989	2000	2000	2000	2007		
Bosnia and Herzegovina	1992						No	Yes
Bulgaria	1981	1986	1989	1998	2005	2005		Yes
Canada	1981	1985	1991		1997			
Croatia	1992		2008	2008	1999	2008	No	Yes
Cyprus	1991		2004		2006	2007		
Czech Republic	1993	1993	1993	1997	1997	2004	No	Yes
Denmark	1982	1986	1993	1996	1997	2002		
Finland	1981	1986	1990	1994	1998	2003		
France	1981	1986	1989	1997	1997	2007		
Georgia	1999							Yes
Germany	1982	1986	1990	1994	1998	2004		Yes
Greece	1983		1998		1998			
Hungary	1980	1986	1991	1995	2002	2006	No	
Iceland	1983							
Ireland	1982		1994		1998			
Italy	1982	1990	1992	1995	1998			
Latvia	1994					2004	No	Yes
Liechtenstein	1983	1986	1994	1994	1997			
Lithuania	1994	2007	2006	2007	2008	2004	No	Yes
Luxembourg	1982	1987	1990	1993	1996	2001		
Macedonia FYR	1997	2010	2010	2010	2010	2010	No	Yes
Malta	1997							
Moldova	1995							Yes
Monaco	1999			2001	2002			
Netherlands	1982	1986	1989	1993	1995	2004		
Norway	1981	1986	1989	1993	1995	2002		
Poland	1985						No	Yes
Portugal	1980					2005		
Romania	1991					2003		
Russian Federation	1980	1986	1989				No	Yes
Slovak Republic	1993	1993	1993	1999	1998	2005	No	Yes
Slovenia	1992		2006		1998	2004	No	Yes
Spain	1982		1990	1994	1997	2005		
Sweden	1981	1986	1990	1993	1995	2002		
Switzerland	1983	1987	1990	1994	1998	2005		
Turkey	1983							
Ukraine	1980	1986	1989				No	Yes
United Kingdom	1982		1990	1994	1996	2005		
United States	1981		1989			2004		

Notes: The list includes countries that have ratified the LRTAP Convention. Source: Mitchell (2015). The years indicate the ratification year of the protocol. The column labeled *GDP* indicate countries where data on GDP is lacking. The column labeled *Structural* indicate countries that have undergone large structural changes in the period analyzed, such as former USSR countries, former Yugoslavia (incl. Albania), former Czechoslovakia, Germany or Poland. Countries lacking data on GDP or undergoing large structural changes are excluded from the analysis. Note that there are five countries that have signed, but not yet ratified the Geneva protocol: Canada, Greece, Portugal, Ukraine, and the United States.

A.2 The links between SO_2, NO_X and VOCs

Sulfur dioxide (SO_2) is part of a group of highly reactive gases known as *oxides of sulfur*. The main source of SO_2 is industrial activity that processes materials containing sulfur, such as electricity generation from coal and oil. Smaller sources of SO_2 emissions include extraction of metal from ore, and the burning of high sulfur fuels by large ships and non-road equipment (EPA, 2015b). Nitrogen oxides (NO_x) is a generic term for the two gases nitric oxide (NO) and nitrogen dioxide (NO_2). The largest share of man-made NO_x emissions are generated by the transportation sector, while a smaller share is emitted from stationary sources such as power generation. Volatile organic compounds (VOCs) are chemicals that easily evaporate, and are emitted by a wide array of products, including paint, cleaning supplies, pesticides, building materials, glues and adhesives, permanent markers, and photographic solutions. Other sources of VOCs are storage and transportation of crude oil on ships leads, and road traffic.

When NO_x and SO_2 emissions reach the atmosphere, they are transformed into acidifying substances. These substances are carried far from their sources by wind. Depending on the weather, the acid pollutants fall to the ground in wet form (acid rain, snow, mist or fog) or in dry form (acid gases or dusts) (EPA, 2015a), causing damage to forest, lakes, rivers, costal waters and man-made structures, such as buildings (OECD, 2014). SO_2, NO_x, and VOCs emissions affect ambient air quality, and are linked to a number of adverse health effects such as respiratory problems, heart disease, and premature mortality. NO_x also reacts with VOCs to form a particularly harmful pollutant: ground-level ozone, which implies that initiatives to combat ozone depletion will potentially target both substances.

As the power sector constitutes a major source of emissions for both SO_2 and NO_x, initiatives to combat either of the substances could potentially have implications for both. Example include plants installing scrubbers on coal-fired unites, improved energy efficiency and conservative initiatives and switching from coal or oil to natural gas. Further, switching from high to low sulfur coal, will reduce SO_2 emissions, while installing low-NO_x burners will reduce NO_x emissions.

A.3 Empirical studies on the effects of LRTAP protocols: a summary

Table A.2: Summary of previous findings

Paper	Protocol	Method	Sample	Period	Findings
Murdoch et al. (1997)	Helsinki, Sofia	Spatial model	25 European countries	1980-1990	Helsinki success, Sofia not
Murdoch et al. (2003)	Helsinki	Spatial model	25 European countries	1980-1990	Incentives to free ride
Finus and Tjøtta (2003)	Oslo	Numerical model	Ratifying countries		No reductions beyond the Nash equilibrium
Ringquist and Kostadinova (2005)	Helsinki	OLS, fixed effects, random effects	19 European countries	1980-1994	No effect
Bratberg et al. (2005)	Sofia	DiD	23 European countries	1985-1996	Significant, but small effect (2.1% annually)
Naughton (2010)	Helsinki, Oslo, Sofia	(2SLS) spatial lag model	16 European countries	1980-2000	Only effect of Sofia
Aakvik and Tjøtta (2011)	Helsinki, Oslo	DiD, country-specific time trends	30 European countries	1960-2002	No effect of Helsinki or Oslo

Appendix B Inference on the pooled estimate

B.1 Inference using the Irwin-Hall distribution

Assuming independent ranks, there is a simple way of arriving at cut-off values: the exact distribution of the mean percentile rank $\bar{p}$ can be calculated using the Irwin-Hall distribution of the sum of E independent uniform random variables. To generate the cut-off values, I do the following. First, I generate E=20 variables labeled u_1.......u_{20}. For each variable u_e, I generate one million uniformly distributed random observations on the interval [0,1]. These can be seen as randomly generated percentile ranks. The distribution of u_1 is shown in Figure B.1a. For E=1 it is straight forward to find the cut-off values: you simply identify the value of $\bar{p}$ in the appropriate percentile. Using one million iterations, the value of $\bar{p}$ corresponding to the 5 percentile should be approximately 0.05. This is the value reported in the first row of Table B.1.

Table B.1: Significance cut-offs assuming independent ranks. Continuous values

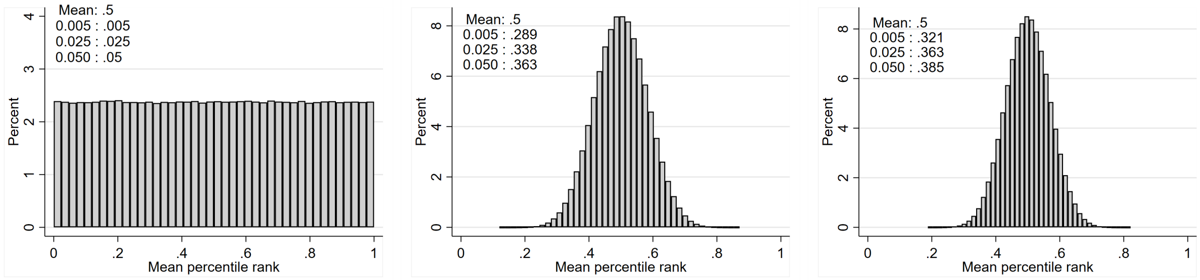
	Percentile					
E	0.5	2.5	5	95	97.5	99.5
1	.005	.025	.050	.950	.975	.995
2	.050	.112	.158	.842	.889	.951
3	.103	.177	.223	.777	.823	.897
4	.148	.220	.262	.738	.780	.853
5	.181	.249	.287	.713	.750	.819
6	.206	.271	.306	.694	.729	.793
7	.227	.287	.320	.680	.712	.773
8	.243	.301	.332	.668	.699	.757
9	.258	.312	.341	.659	.688	.742
10	.269	.322	.350	.650	.678	.731
11	.280	.330	.357	.644	.670	.721
12	.289	.338	.363	.637	.663	.711
13	.297	.344	.368	.632	.657	.703
14	.304	.349	.373	.627	.651	.696
15	.310	.354	.377	.623	.646	.689
16	.316	.359	.381	.619	.641	.683
17	.321	.363	.385	.615	.637	.678
18	.327	.367	.388	.612	.633	.673
19	.331	.370	.391	.609	.630	.669
20	.335	.374	.394	.606	.626	.665

Notes: Simulated using one million iterations of the mean of E uniformly distributed variables on [0,1].

For E=2, I start by taking the mean of u_1 and u_2. This can be seen as taking the mean of

Figure B.1: Distribution of simulated percentile ranks. Continuous values

(a) E=1 **(b)** E=12 **(c)** E=17



Notes: Figures show the distribution for the mean of E uniformly distributed variables on [0,1]. Variables are simulated using one million iterations

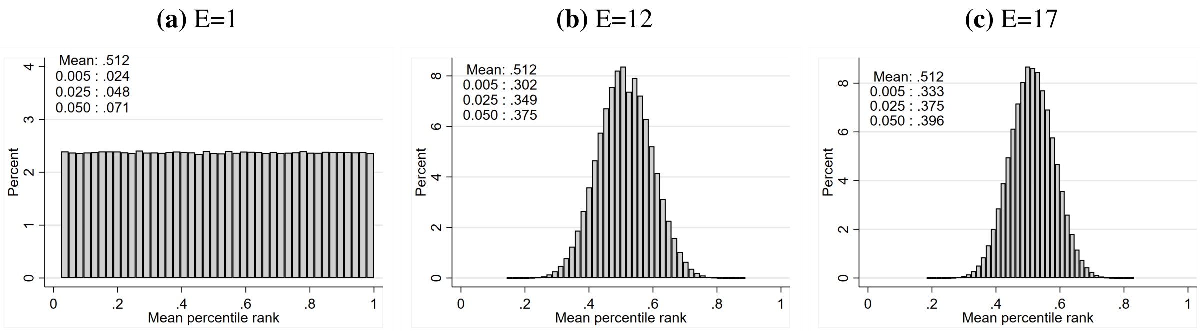
percentile ranks for two treated countries. As each of the variables u_1 and u_2 has one million observations, the new variable also has one million observations. To find cut-off values, I identify the value corresponding to the chosen percentile. For the 5 percentile, the value will be larger compared to the case with E=1. This comes from the fact that averaging over several variables makes it less likely to arrive at extremely large and extremely small values. The cut-off values for E=2 are reported in the second row of Table B.1.

The procedure is then repeated for values E=3,...,20. The distributions of the mean percentile ranks when E=12 and E=17 are shown in Figures B.1b and Figure B.1c. Table B.1 shows the percentiles of the distribution for E=1,...,20 treated countries. For 12 treated countries, a two-sided 5% significance test requires the mean percentile rank $\bar{p}$ to be below 0.338 or above 0.663. For 17 treated countries, a two-sided 5% significance test requires the mean percentile rank $\bar{p}$ to be below 0.363 or above 0.637.

B.2 Inference using the Irwin-Hall distribution with discrete values

The cut-off values listed in Table B.1 are based on the mean of E uniformly distributed variables on [0,1]. We interpreted these variables as percentile ranks. The country-specific percentile ranks, however, will never be smaller than $\frac{1}{42}$. This comes from trimming the pool of countries down to 42 (1 treated and 41 donors). Further, the distribution will be discrete as there is a limited number of possible observations of p_j. As a result, the distribution of mean percentile ranks will be slightly skewed to the right.

Figure B.2: Distribution of simulated mean percentile ranks. Discrete values



Notes: Figures show the distribution for the mean of E uniformly distributed variables on [$\frac{1}{42}$,1]. Discrete values. Variables are simulated using one million iterations.

Table B.2: Significance cut-offs assuming independent ranks. Discrete values.

	Percentile					
E	0.5	2.5	5	95	97.5	99.5
1	.024	.048	.071	.952	.976	1.000
2	.060	.119	.167	.857	.905	.964
3	.119	.190	.238	.786	.833	.905
4	.161	.232	.274	.750	.792	.863
5	.190	.262	.300	.724	.762	.833
6	.218	.282	.317	.706	.742	.806
7	.238	.299	.333	.690	.724	.786
8	.256	.313	.345	.682	.711	.768
9	.270	.325	.354	.669	.698	.757
10	.281	.333	.362	.662	.690	.743
11	.292	.342	.368	.656	.682	.732
12	.302	.349	.375	.649	.675	.724
13	.310	.355	.381	.645	.668	.716
14	.316	.362	.386	.639	.663	.709
15	.322	.367	.389	.635	.657	.703
16	.329	.372	.393	.631	.653	.696
17	.333	.375	.396	.627	.648	.690
18	.339	.380	.401	.624	.646	.687
19	.343	.382	.404	.622	.642	.682
20	.348	.386	.406	.618	.638	.677

Notes: Simulated using one million iterations of the mean of E uniformly distributed variables on [$\frac{1}{42}$,1]. Discrete values.

Figure B.2 shows the distribution of the mean percentile rank when simulated observations are constrained to take on discrete values on the interval [$\frac{1}{42}$,1]. Table B.2 shows the corresponding cut-off values. These cut-off-values will be more appropriate for determining statistical significance in our case. For 12 treated countries, a two-sided 5% significance test requires the mean percentile rank $\bar{p}$ to be below 0.349 or above 0.675. For 17 countries, a two-sided 5%

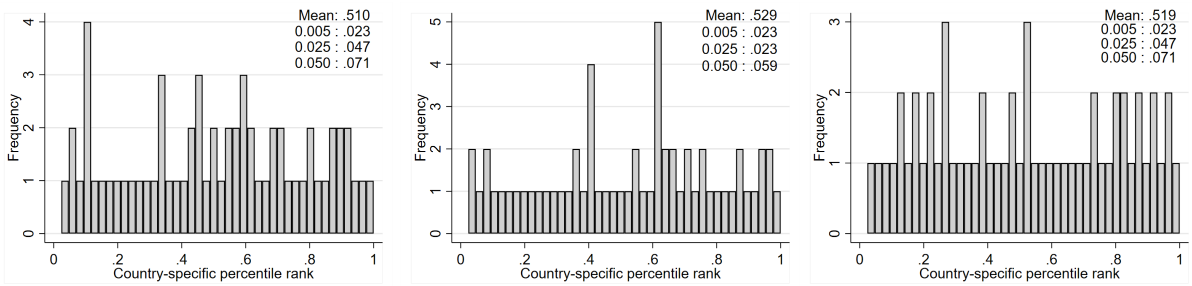
significance test requires the mean percentile rank $\bar{p}$ to be below 0.375 or above 0.648.

B.3 Inference using a permutation procedure

Assuming independent ranks, I can also perform inference on the pooled treatment estimate by randomly permuting the treatment status.[72] By iterating the permutation procedure one million times, I get a distribution of the mean percentile ranks. To conduct inference using a permutation procedure, I do the following. I start by applying the synthetic control method and placebo-based inference for each of the countries in the dataset – both actually treated (E) and donor countries ($N-E$). Next, I construct a dataset containing the percentile ranks for actually treated countries and donor countries. In the case of the Helsinki protocol, the dataset includes 12 treated countries and 51 donor countries, which totals to 63 countries. The percentile ranks in this dataset will take on values between $\frac{1}{42}$ and 1.[73] The distribution of the country-specific ranks for the three protocols are shown in Figure B.3.

Figure B.3: Distribution of country-specific percentile ranks.

(a) Helsinki (SO_2) **(b)** Sofia (NO_x) **(c)** Geneva (VOCs)



Notes: Figures show histograms of country-specific percentile ranks in the case of the Helsinki (a), Sofia (b) or Geneva (c) protocol. The vertical axis denotes the number of countries in each of the 42 bins. 0.005 denotes the 0.5 percentile, 0.025 denotes the 2.5 percentile and 0.05 denotes the 5 percentile. For Helsinki, the total number of countries N is 63, for Sofia N=60 and for Geneva N=57.

Next, I randomly permute the treatment status for countries in the dataset. In other words: I randomly select E countries from the dataset, which are labeled as "treated". For the Helsinki protocol, this means that I pick 12 countries from a pool of 63 countries. I then calculate the average of the country-specific percentile ranks for the E randomly selected countries. This

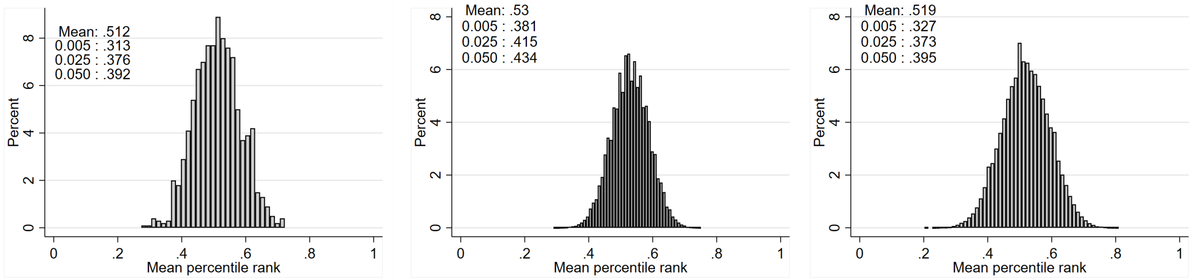
[72] The procedure has similarities to procedures described in Section 4.5 in Dube and Zipperer (2015) and in Gobillon and Magnac (2016).

[73] The number 42 comes from the procedure of trimming down the dataset to 42 countries before using the synthetic control method. Including 42 countries allows me to evaluate significance at a 5% level using a two-sided test in the case of a single treated country.

mean percentile rank is added to a new dataset. The procedure is iterated one million times, resulting in a distribution of mean percentile ranks $\bar{p}$. The distributions of the mean percentile ranks for each of the three protocols are presented in Figure B.4.

Figure B.4: Distribution of the mean percentile rank. Randomization procedure assuming independent ranks.

(a) Helsinki (SO_2) **(b)** Sofia (NO_x) **(c)** Geneva (VOCs)



Notes: Figures show histograms of the mean percentile rank using the randomization procedure described in Section B.3. Values are generated using one million iterations.

To find the cut-off values for e.g., a two-sided 1% significance level, I identify the value of the mean percentile rank $\bar{p}$ corresponding to the 0.5 and 99.5 percentile. To find the cut-off values corresponding to a 5% significance level, I identify the value of the mean percentile rank $\bar{p}$ corresponding the 2.5 and 97.5 percentile, etc. The cut-off values corresponding to a 1, 5 and 10% significance level for the three different interventions (*Helsinki, Sofia, Geneva*) are presented in Table B.3, and marked by *Independent ranks (randomization)*.

The cut-off values should be close to those generated using the Irwin-Hall distribution, but they will not necessarily be identical. There are several reasons for this. First, while the individual rank-based p-values calculated for each intervention will be uniformly distributed on the interval $[\frac{1}{42}, 1]$, the percentile ranks in the dataset of N countries (63 in the case of Helsinki) will not necessarily be uniformly distributed on the interval $[\frac{1}{42}, 1]$. This can be seen from Figure B.3.[74] As a result, the distribution of mean percentile ranks might be slightly different to the ones in Table B.2. Second, while the randomization procedure simultaneously picks E countries from the pool of N countries, the Irwin-Hall simulation procedure can be seen as

[74] This is caused by three things. First, as N is a relatively small number (63, 60 or 57), the values in the sample will not be exactly uniformly distributed. When $N \to \infty$, the distribution will be approximately uniform. Second, when I run the synthetic control method for each country, I trim the sample down to 42 countries. Each "treated" country will therefore have a different set of countries in the pool of 41 donors. Third, when applying the synthetic control method for actually treated countries, I exclude other actually treated countries from the donor pool. In theory, all actually treated countries could therefore be assigned a rank-based p-value of $\frac{1}{42}$.

Table B.3: Comparing significance cut-offs from different procedures

Protocol	Assumption (procedure)	E	Percentile 0.5	2.5	5	95	97.5	99.5
Helsinki (SO_2)	Independent ranks (Irwin-Hall, continuous)	12	.289	.338	.363	.637	.663	.711
Helsinki (SO_2)	Independent ranks (Irwin-Hall, discrete)*	12	.302	.349	.375	.649	.675	.724
Helsinki (SO_2)	Independent ranks (randomization)	12	.313	.376	.392	.627	.652	.700
Helsinki (SO_2)	Dependent ranks (randomization)	12	.278	.284	.300	.792	.792	.792
Sofia (NO_x)	Independent ranks (Irwin-Hall, continuous)	17	.321	.363	.385	.615	.637	.678
Sofia (NO_x)	Independent ranks (Irwin-Hall, discrete)*	17	.333	.375	.396	.627	.648	.690
Sofia (NO_x)	Independent ranks (randomization)	17	.381	.415	.434	.626	.642	.676
Sofia (NO_x)	Dependent ranks (randomization)	17	.399	.417	.437	.707	.707	.707
Geneva (VOCs)	Independent ranks (Irwin-Hall, continuous)	12	.289	.338	.363	.637	.663	.711
Geneva (VOCs)	Independent ranks (Irwin-Hall, discrete)*	12	.302	.349	.375	.649	.675	.724
Geneva (VOCs)	Independent ranks (randomization)	12	.327	.373	.395	.642	.664	.708
Geneva (VOCs)	Dependent ranks (randomization)	12	.290	.290	.290	.764	.764	.764

Notes: Table shows cut-off values for $\bar{p}$ using different inference procedures. E equals the number of treated countries. With a two-sided test, a 5% significance level equals a value below the 2.5 percentile or above the 97.5 percentile. Values are generated using one million iterations. The Irwin-Hall cut-off values are the ones generated using discrete values on the interval $[\frac{1}{42},1]$ The preferred procedure is indicated by *.

drawing E countries with replacement. The Irwin-Hall procedure could therefore, in theory, pick E percentile ranks with value $\frac{1}{42}$. For the randomization procedure, however, the number of percentile ranks with value $\frac{1}{42}$ is constrained by the number of actually observed percentile ranks in the dataset. From Figure B.3 we see that there is only one country with a percentile rank of $\frac{1}{42}$ in the case of the Helsinki protocol.

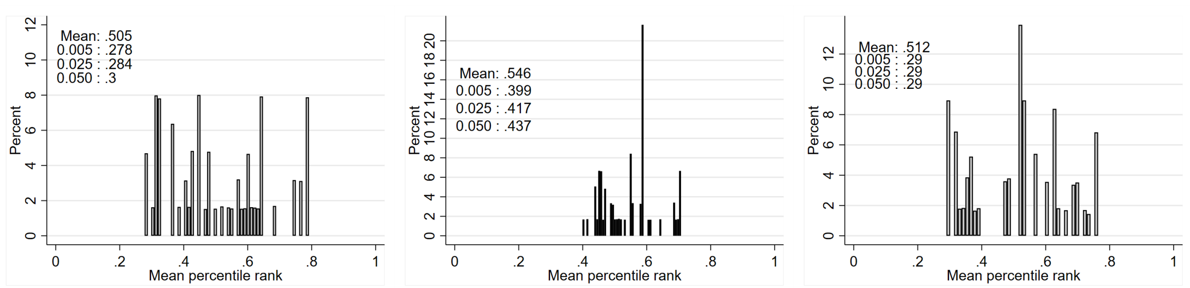
When we compare the cut-off values in Table B.3 to the ones generated from the Irwin-Hall distribution (Table B.2), the new cut-off values are somewhat higher. In the case of the Helsinki protocol, where we have 12 treated countries, the mean percentile rank $\bar{p}$ needs to be below 0.376 in order for the treatment effect to be significant at a 5% level (using a two-sided test). The corresponding value using the Irwin-Hall distribution is 0.349. For Sofia, we have 17 treated countries, resulting in a higher cut-off value for a 5% significance level (0.415). The corresponding value using the Irwin-Hall distribution is 0.375. The Geneva protocol has 12 treated countries, and the mean percentile rank $\bar{p}$ needs to be below 0.373 in order for the treatment effect to be significant at a 5% level.

B.4 Permutation procedure that accounts for donor overlap

A potential problem with the procedures presented in the previous sections is that they do not account for rank dependency. In the analysis, treated countries will often be assigned many of the same donors, resulting in a dependency across the percentile ranks. The inference procedure described in Section B.3 randomly picks countries as "treated", which means that the procedure does not account for potential donor overlap in the analysis.

Figure B.5: Distribution of the mean percentile rank. Randomization procedure accounting for donor overlap.

(a) Helsinki (SO_2) **(b)** Sofia (NO_x) **(c)** Geneva (VOCs)



Notes: Figures show histograms of the mean percentile rank using the procedure described in Section B.4. Values are generated using 10 000 iterations.

To try to address the problem of rank dependency, I repeat the permutation procedure described in Section B.3, but use a procedure to assign treatment status that accounts for potential correlation between treated countries. Specifically, I start by randomly selecting one country from the pool of N countries to be "treated". Next, I find the $(E-1)$ countries with the smallest geographical distance to the "treated" country by using latitude and longitude of the countries' centroids. This accounts for the feature that actually treated countries are geographically clustered. I then take the average of the percentile ranks for the E "treated" countries, and add this mean value to a new dataset.

The rest of the procedure is similar to the one described in Section B.3. The distribution of mean percentile ranks is shown in Figure B.5. The corresponding cut-off values are listed in Table B.3, marked by *Dependent ranks (randomization)*. Accounting for rank-dependency gives lower cut-off values, meaning that it is harder to reject the null hypothesis. Due to the few possible combinations of treated countries under the geographical location constraint, I cannot identify unique cut-off values for the 0.5, 2.5 and 5 percentile in the case of the Geneva protocol.

Appendix C Data and descriptive statistics

Table C.1: Country acronyms

Acronym	**Country name**
AUT	Austria
BEL	Belgium
CAN	Canada
CHE	Switzerland
DNK	Denmark
ESP	Spain
FIN	Finland
FRA	France
GBR	England
GRC	Greece
IRL	Ireland
ITA	Italia
LUX	Luxemburg
NLD	The Netherlands
NOR	Norway
SWE	Sweden
USA	United States

Table C.2: Trimming the sample

Step	Category	Adjustments
Before analysis		
1	Large restructuring	Drop former USSR countries, former Yugoslavia (incl. Albania), former Czechoslovakia, Germany, Poland, Bulgaria, Romania, Mongolia. Drop countries with long-lasting wars during the time period. Drop Norway when analyzing VOCs due to accelerating oil production.
2	Data availability	Drop countries with missing values on emissions, GDP, or fossil fuel share.
3	Similarity	Drop poorest 20% and richest 1% (in 1980). Drop countries with extremely high or low pollution levels (SO_2 per capita ≤ 1, VOC per capita ≥ 600). Drop small island states and microstates.
4	Volatility	Drop countries if the maximum (normalized) emissions in the time series is 3 times higher than the lowest (normalized) emissions in the time series. Drop countries with extreme treatment effects (Jordan (SO_2), Korea Rep (NO_x), Israel (VOC))
5	Spillovers	Drop LRTAP countries that have signed but not ratified the protocol in question. Drop Iceland when estimating the effects of Sofia and Geneva due to the adoption of a similar policy
Before inference		
6	Pre-treatment fit	Trim donor pool down to 41 control countries (+ 1 treated) based on the MSPE and the pre-treatment, per capita pollution level.

Table C.3: Treated and donor countries.

	Panel A: Helsinki (SO_2)
Treated: 12	Austria, Belgium, Canada, Denmark, Finland, France, Italy, Luxembourg, Netherlands, Norway, Sweden, Switzerland
Donor: 51	Algeria, Argentina, Australia, Bahrain, Bolivia, Brazil, Cameroon, Chile, Colombia, Costa Rica, Cote dIvoire, Cuba, Cyprus, Ecuador, Egypt, El Salvador, Greece, Honduras, Hong Kong SAR, Iceland, Iran Rep, Ireland, Israel, Japan, Korea Rep, Malaysia, Malta, Mexico, Morocco, New Zealand, Nicaragua, Nigeria, Panama, Paraguay, Peru, Philippines, Portugal, Saudi Arabia, Senegal, Singapore, South Africa, Spain, Syrian Arab Republic, Thailand, Tunisia, Turkey, United Arab Emirates, United Kingdom, United States, Venezuela, Zimbabwe
	Panel B: Sofia (NO_x)
Treated: 17	Austria, Belgium, Canada, Denmark, Finland, France, Greece, Ireland, Italy, Luxembourg, Netherlands, Norway, Spain, Sweden, Switzerland, United Kingdom, United States
Donor: 43	Algeria, Argentina, Australia, Bahrain, Brazil, Cameroon, Chile, Colombia, Costa Rica, Cuba, Cyprus, Ecuador, Egypt, El Salvador, Honduras, Hong Kong SAR, Iran Rep, Israel, Japan, Jordan, Malaysia, Malta, Mexico, Morocco, New Zealand, Nicaragua, Nigeria, Oman, Panama, Peru, Philippines, Portugal, Saudi Arabia, Senegal, Singapore, South Africa, Syrian Arab Republic, Thailand, Tunisia, Turkey, United Arab Emirates, Uruguay, Venezuela
	Panel C: Geneva (VOC)
Treated: 12	Austria, Belgium, Denmark, Finland, France, Italy, Luxembourg, Netherlands, Spain, Sweden, Switzerland, United Kingdom
Donor: 45	Algeria, Argentina, Australia, Bahrain, Brazil, Cameroon, Chile, Colombia, Costa Rica, Cote dIvoire, Cuba, Cyprus, Ecuador, Egypt, El Salvador, Honduras, Hong Kong SAR, Iran Rep, Ireland, Japan, Jordan, Malaysia, Malta, Mexico, Morocco, New Zealand, Nicaragua, Nigeria, Oman, Panama, Paraguay, Peru, Philippines, Saudi Arabia, Senegal, Singapore, South Africa, Syrian Arab Republic, Thailand, Tunisia, Turkey, Uruguay, Venezuela, Zambia, Zimbabwe

Table C.4: Summary of demographics, by treated and donor countries. 1980

	Treated countries		Donor pool		Diff	
	mean	(sd)	mean	(sd)	mean	(se)
SO2 per capita	74	(33)	35	(38)	39	(11)***
NOx per capita	48	(27)	22	(27)	26	(8)***
VOC per capita	55	(14)	54	(62)	1	(10)
GDP per capita (constant 2005 USD)	27,788	(6,609)	8,911	(13,216)	18,877	(2,658)***
Fossil/energy	82	(14)	74	(26)	8	(5)
GDP growth (pct)	2.36	(1.73)	3.33	(6.59)	-0.98	(1.05)
Population growth (pct)	0.38	(0.35)	2.37	(1.52)	-1.99	(0.24)***
Number of countries	12		51			

Notes: Table shows means and standard deviations. The sample corresponds to the sample used for estimating the effect of the Helsinki protocol on (normalized) SO_2 emissions. The two last columns show the difference in means and the standard errors from a t-test on the equality of means. Statistics for NO_x and VOCs are from the years 1987 and 1990, respectively. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix D Annual estimates, weights, and predictor match

D.1 Annual treatment effects

Table D.1: Average treatment effects, by year. SO_2. 1980-1994

Time		Treatment effects			95% CI	
Year	Relative	Mean	HL	P-value	low	high
1980	1	-2.011	-0.667	0.466	-6.500	2.250
1981	2	-12.969	-12.167	0.329	-20.333	0.167
1982	3	-14.772	-8.750	0.359	-21.833	-0.417
1983	4	-20.499	-16.500	0.268	-30.000	-10.750
1984	5	-22.603	-20.000	0.236	-27.083	-11.417
1985	6	-16.857	-13.417	0.355	-22.500	-3.167
1986	7	-18.210	-15.583	0.290	-22.667	-7.583
1987	8	-16.691	-13.833	0.315	-22.583	-7.500
1988	9	-22.810	-15.000	0.302	-27.833	-5.500
1989	10	-21.543	-18.333	0.288	-30.833	-6.167
1990	11	-21.444	-17.333	0.306	-31.750	-4.000
1991	12	-16.489	-9.750	0.387	-29.750	2.417
1992	13	-21.706	-19.333	0.357	-35.167	1.750
1993	14	-23.883	-21.083	0.333	-36.917	-5.417
1994	15	-23.076	-20.583	0.313	-37.750	-9.250

Notes: Critical values for the mean percentile rank are derived from the simulation procedure described in Appendix B.2. Inverting this rank gives the 95% CIs. Main results are reported in Table 3. Significance levels for SO_2: 1%: .302 , 5%: .349 , 10%: .375. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.2: Average treatment effects, by year. NO_x. 1987-1999

Time		Treatment effects			95% CI	
Year	Relative	Mean	HL	P-value	low	high
1987	1	2.221	1.412	0.612	0.235	2.412
1988	2	-3.012	-5.412	0.359	-9.176	-2.000
1989	3	-1.829	-1.235	0.466	-8.765	4.118
1990	4	-6.412	-6.294	0.373	-14.176	-1.471
1991	5	-10.966	-10.882	0.301	-18.824	-6.000
1992	6	-14.652	-13.294	0.287	-17.882	-7.000
1993	7	-17.560	-20.412	0.266	-24.235	-10.471
1994	8	-17.117	-13.353	0.326	-30.235	-1.588
1995	9	-18.026	-12.059	0.321	-28.471	-4.294
1996	10	-17.570	-16.706	0.314	-25.824	-6.235
1997	11	-17.583	-10.176	0.354	-25.471	-3.118
1998	12	-19.014	-16.882	0.333	-33.176	-5.941
1999	13	-25.636	-20.471	0.301	-30.588	-7.059

Notes: Critical values for the mean percentile rank are derived from the simulation procedure described in Appendix B.2. Inverting this rank gives the 95% CIs. Main results are reported in Table 5. Significance levels for NO_x: 1%: .333 , 5%: .375 , 10%: .396. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.3: Average treatment effects, by year. VOC. 1990-1999

Time		Treatment effects			95% CI	
Year	Relative	Mean	HL	P-value	low	high
1990	1	-2.822	-1.750	0.419	-3.083	1.083
1991	2	-3.798	-2.667	0.423	-5.833	2.417
1992	3	-7.253	-4.750	0.381	-9.167	0.333
1993	4	-11.432	-10.583	0.260	-17.750	-5.917
1994	5	-14.866	-19.417	0.266	-27.417	-3.000
1995	6	-15.817	-14.333	0.264	-28.250	-7.083
1996	7	-16.792	-17.083	0.224	-25.250	-9.750
1997	8	-17.450	-18.167	0.210	-24.583	-10.333
1998	9	-16.613	-15.917	0.298	-25.083	-5.667
1999	10	-19.599	-17.083	0.262	-28.083	-10.167

Notes: Critical values for the mean percentile rank are derived from the simulation procedure described in Appendix B.2. Inverting this rank gives the 95% CIs. Main results are reported in Table 7. Significance levels for VOCs: 1%: .302 , 5%: .349 , 10%: .375. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.2 Weights used for constructing synthetic controls

Table D.4: Weights used for constructing synthetic controls. SO_2. 1980.

	AUT	BEL	CAN	CHE	DNK	FIN	FRA	ITA	LUX	NLD	NOR	SWE
Algeria						0.2	0.1	0.1				
Argentina						0.3	0.2	0.1				
Australia						1.3	0.3	0.2				
Bahrain	38.0					0.4		45.2		34.9		
Brazil						0.3	0.1	0.1				
Chile			28.1			18.1						7.4
Colombia						0.3	0.1	0.1				
CostaRica						0.3	0.1	0.1				
CotedIvoire						0.3	0.1	0.1				
Cuba						0.4	0.1	0.1				
Cyprus						0.5	0.2	0.2				
Ecuador						0.2	0.1	0.1				
ElSalvador						0.3	0.1	0.1				0.3
Greece						0.6	0.2	0.3				
Honduras						0.3	0.1	0.1				
HongKongSAR						0.5	0.2	0.2				
Iceland	25.4	0.4			9.1	28.0	15.3	3.4	1.2	36.4	78.3	45.4
IranRep						0.3	0.1	0.1				
Ireland						0.9	0.2	0.2				
Israel						0.6	0.2	0.2				
Japan	3.4					0.6	38.2	6.1				
KoreaRep						0.3	0.1	0.1				
Malaysia						0.3	0.1	0.1				
Malta						0.4	0.1	0.1				
Mexico						0.4	0.1	0.1				
NewZealand		7.6	32.4	85.6	19.4	0.5	1.2	0.2	43.2		8.6	23.2
Nicaragua						0.3	0.1	0.1				
Panama						0.3	0.1	0.1				
Paraguay						0.3	0.1	0.1				
Peru						0.4	0.1	0.1				
Portugal						0.4	0.1	0.1				
SaudiArabia						0.4	0.2	2.2		15.6		
Singapore						0.5	0.2	0.2				
SouthAfrica						0.5	0.1	0.1				
Spain						0.7	0.2	0.2				
SyrianArabRepublic						0.2	0.1	0.1				
Tunisia						0.3	0.1	0.1				
Turkey						0.3	0.1	0.1				
UnitedArabEmirates				13.5								
UnitedKingdom	33.2	50.0				0.9	31.4	8.2				
UnitedStates		42.1	39.6	1.0	71.6	37.5	9.3	31.0	55.6	13.1	13.1	23.5
Venezuela						0.3	0.1	0.1				

Table D.5: Weights used for constructing synthetic controls. NO_x. 1987.

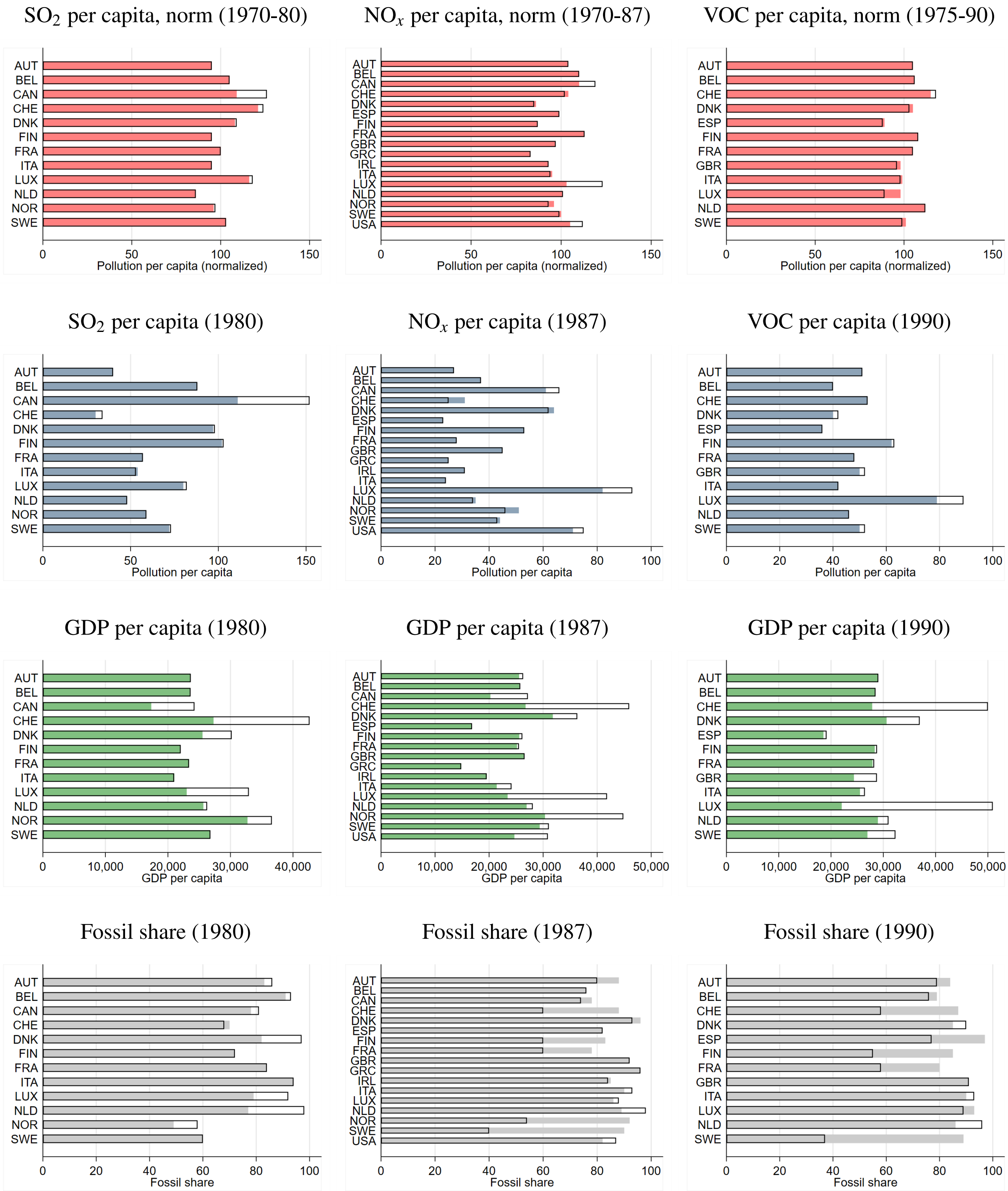
	AUT	BEL	CAN	CHE	DNK	ESP	FIN	FRA	GBR	GRC	IRL	ITA	LUX	NLD	NOR	SWE	USA
Algeria						1.0			0.1	0.4	0.2						
Argentina						0.6			0.9	0.3	0.2						
Australia			26.4			0.7				4.4	5.2	3.6	54.6				22.7
Bahrain						2.1			0.1	1.9	1.2						
Brazil						1.0				0.3	0.2						
Cameroon		13.0				0.5		7.6		0.2	0.1						
Chile						0.9				0.3	0.2						
Colombia						0.9				0.3	0.2						
CostaRica						0.9				0.3	0.2						
Cuba						1.0				0.3	0.2						
Cyprus						2.2			15.3	54.4	0.7						
Ecuador						1.1				0.3	0.2						
Egypt						1.0				0.3	0.2						
ElSalvador		2.7				1.0		4.6		0.2	0.2						
Honduras		0.4				1.2	4.0			0.2	0.2						
HongKongSAR	20.0			16.0	31.2	8.6	23.4		1.4	19.8	29.9	45.1		21.4	19.8	16.4	
IranRep						0.8			0.4	0.4	0.2						
Israel						1.8			0.2	0.8	0.6						
Japan	73.2	64.6		72.8	21.8	38.6		78.4	41.9	6.4	24.1	51.2		63.5	47.7	58.6	
Jordan						1.1			0.1	0.4	0.2						
Malaysia						0.8			0.1	0.3	0.2						
Malta						1.7				0.5	0.5						
Mexico						1.2			0.1	0.4	0.3						
Morocco						0.9			0.1	0.3	0.2						
NewZealand		0.1				12.1	45.5		0.1	0.6	30.5						
Nicaragua						0.9				0.2	0.2						
Oman						1.4			9.7	0.8	0.4						
Panama						0.7				0.2	0.2						
Peru						0.7				0.2	0.2						
Philippines		0.1				0.5				0.2	0.1						
Portugal						2.3			0.1	0.6	0.6						
SaudiArabia						1.6			0.1	0.5	0.6						
Senegal						0.7				0.2	0.1						
Singapore						1.9			0.2	1.1	0.7						
SouthAfrica						0.5			2.3	0.3	0.2						
SyrianArabRepublic						0.9			0.1	0.3	0.2						
Thailand						1.1				0.3	0.2						
Tunisia						0.9				0.3	0.2						
Turkey						1.1				0.3	0.2						
UnitedArabEmirates	6.8	18.7	27.4	11.1	47.0		27.0	9.4	26.1				22.0	15.1	32.5	25.1	39.9
Uruguay			46.2			0.4				0.2	0.1		23.5				37.4
Venezuela						0.8			0.3	0.3	0.2						

Table D.6: Weights used for constructing synthetic controls. VOCs. 1990.

	AUT	BEL	CHE	DNK	ESP	FIN	FRA	GBR	ITA	LUX	NLD	SWE
Australia		1.1	6.3			15.3					3.5	
Bahrain	8.0	0.4		3.3	10.5	5.8	4.7	12.1	7.8	26.5		10.3
HongKongSAR	0.6				80.0			35.2	31.7	34.8		18.5
Japan	83.1	89.9	78.3	96.7	9.5	75.8	83.2	52.7	60.4	38.7	85.9	71.2
NewZealand	8.3						5.8					
Paraguay		8.7				3.1	6.3					
SaudiArabia			15.4								10.6	

D.3 Match of predictors

Figure D.1: Predictor match (Helsinki, Sofia, Geneva).



Notes: Hollow bars indicate treated countries. Colored bars indicate synthetic controls.. First column report results from the Helsinki protocol. Second column report results from the Sofia protocol. Third column report results from the Geneva protocol.

Appendix E The Helsinki Protocol: comparing results to previous findings

In this Appendix, I use a difference-in-difference (DiD) approach to estimate the effects of the Helsinki protocol. The purpose is twofold. First, by applying a similar strategy as in Aakvik and Tjøtta (2011), I can identify potential reasons for why they arrive at a small and insignificant treatment effect. Is it the methodology, the sample, or the chosen treatment year? Second, using a DiD approach may help illustrate potential (dis)advantages of a synthetic control method.

E.1 Methodology: difference-in-differences

When $Helsinki_{jt}$ is a dummy that indicates if country j is treatment at time t, the DiD estimator can be written as:

$$y_{jt} = \beta_1 Helsinki_{jt} + \gamma' \mathbf{X}_{jt} + c_j + \delta_t + \varepsilon_{jt}, \tag{10}$$

where j is country, t is time (year), y_{jt} is country-level emissions, $\mathbf{X}_{jt}$ is a vector of observable covariates, c_j are country-specific fixed effects, δ_t are time dummies and ε_{jt} is the idiosyncratic error term.[75] The DiD set-up relies on the assumption that the treatment group and the control group would have followed parallel trends in absence of treatment. While this is an untestable assumption, comparing the pre-intervention trend indicates if the assumption holds or not. To verify if the pre-treatment trends are parallel, I include leads and lags dummies indicating years relative to the intervention. If we denote M as the number of leads and K as the number of lags, we can estimate the unfolding of the treatment with the regression:

$$y_{jt} = \sum_{m=0}^{M} \beta_{-m} Helsinki_{jt-m} + \sum_{k=1}^{K} \beta_{+k} Helsinki_{jt+k} + \gamma' \mathbf{X}_{jt} + c_j + \delta_t + \varepsilon_{jt}, \tag{11}$$

where lead m captures potential deviations in the pre-treatment m years before the intervention, and lag k captures the effect of the treatment k years after the intervention. The estimated coefficients for leads dummies (β_{-m}) should show no effect of treatment under the parallel trends assumption, while the coefficients for the lags dummies (β_{+k}) capture how the treatment

[75]For the DiD to give consistent estimates, we need to assume that the error term is not correlated with time-varying omitted variables: $E[\varepsilon_{jt}|c_j, \mathbf{X}_{j1}, \ldots\ldots, \mathbf{X}_{jT}] = 0$.

effect unfolds over time.

Aakvik and Tjøtta (2011) addresses the potential problem of different trends by including country-specific (linear and quadratic) trends in the DiD estimation. While including such trends may help mitigate potential problems of nonparallel trends, it may also absorb large parts of the treatment effect. It also changes the interpretation of the DiD estimates: the estimated treatment effects will now reflect the deviations in emissions from the country-specific, linear trend. If we denote t as a linear time trend, we can estimate the unfolding of the treatment with the regression:

$$y_{jt} = \sum_{m=0}^{M} \beta_{-m} Helsinki_{jt-m} + \sum_{k=1}^{K} \beta_{+k} Helsinki_{jt+k} + \gamma' \mathbf{X}_{jt} + c_j + \delta_t + h_j t + \varepsilon_{jt}, \tag{12}$$

where $h_j t$ are the country-specific, linear trends. The estimated coefficients for lags dummies (β_{+k}) will now reflect the difference between the treatment group and the control group in the deviation in emission from the country-specific linear trend.

E.2 Results

Figure E.1 shows the results from the DiD estimation. Results are shown for three different samples, indicated by the column heading.[76] The first row plots the (raw) average of log SO_2, by treatment and control group, while row 2 plots the coefficients β_{-m} and β_{-m} estimated from equation 11. Although the level and development in log SO_2 for the control group varies across the samples (see first row), the estimated treatment effect is very similar (see second row). This implies that the definition of the control group has limited effect on the treatment effect.

Further, we see that the pre-treatment trend is significantly different for the treatment and control group.[77] Using 1985 as the treatment year, instead of 1980, the pre-treatment trend is clearly not parallel, and the estimated treatment effect is lower (see third row).

Rows 4 and 5 plot the coefficients β_{-m} and β_{-m} estimated from equation 12. Including

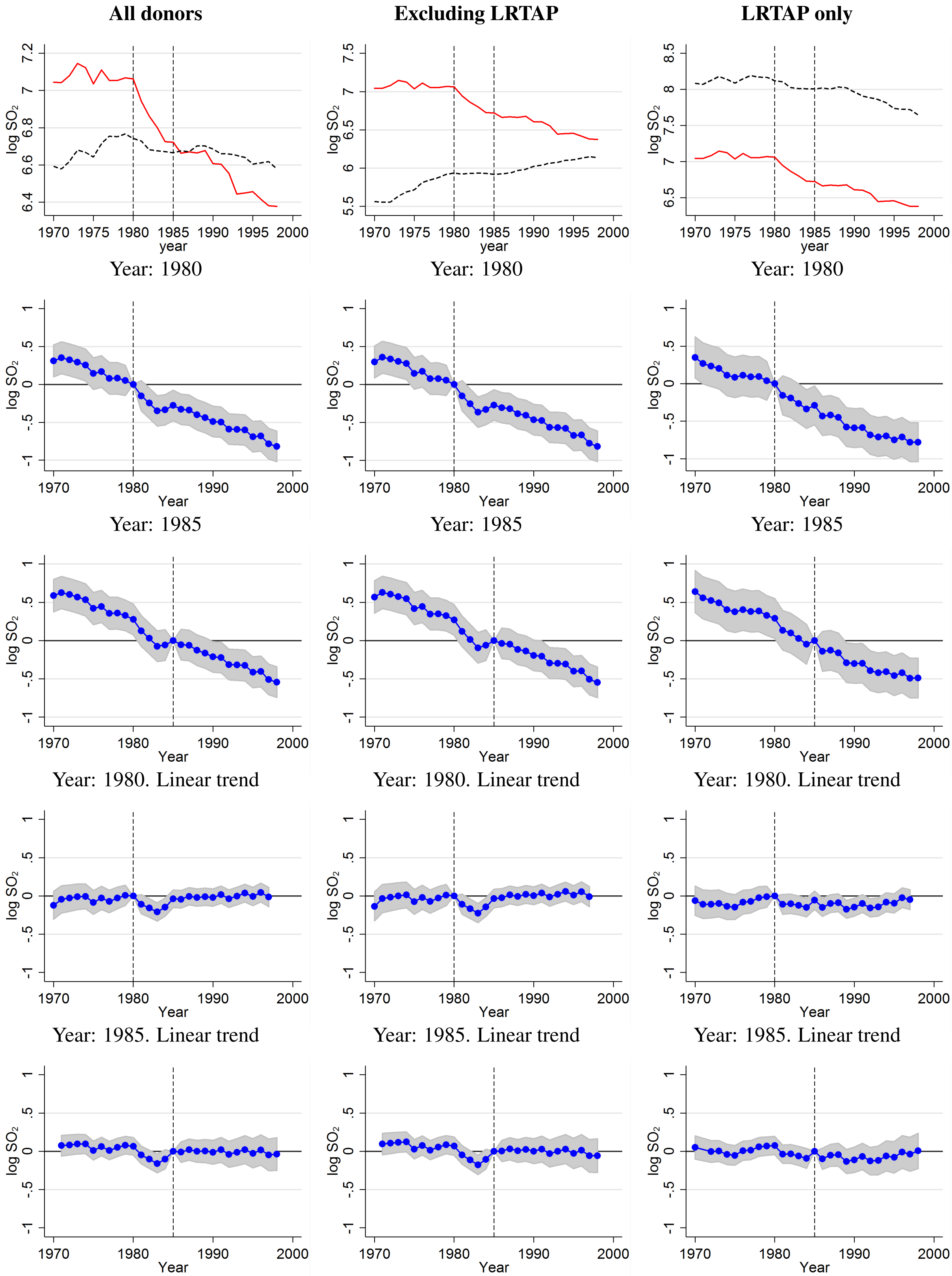
[76]The first column (*All donors*) uses all countries in the donor pool as control countries. The second column (*Excluding LRTAP*) excludes other LRTAP countries from the control group. The third column (*LRTAP only*) restricts the control group to LRTAP countries only.

[77]While is does not seem to be significantly different for the last sample (*LRTAP only*), this is partly due to wider confidence intervals.

linear trends seem to absorb almost all of the treatment effect.

Taken together, I find that the definition of the control group has minor effects of the treatment effect, using 1985 as the treatment year lowers the effect, and including a linear trend wipes out the treatment effect completely - with the exception of a few years in some of the specifications. The results imply that we clearly need to address the parallel trend assumption, but that including linear trends may lead to an underestimation of the treatment effect. Using a synthetic control method offers an alternative approach to the problem of different trends.

Figure E.1: Difference-in-differences. Log SO_2. Different **treatment years** and **samples**

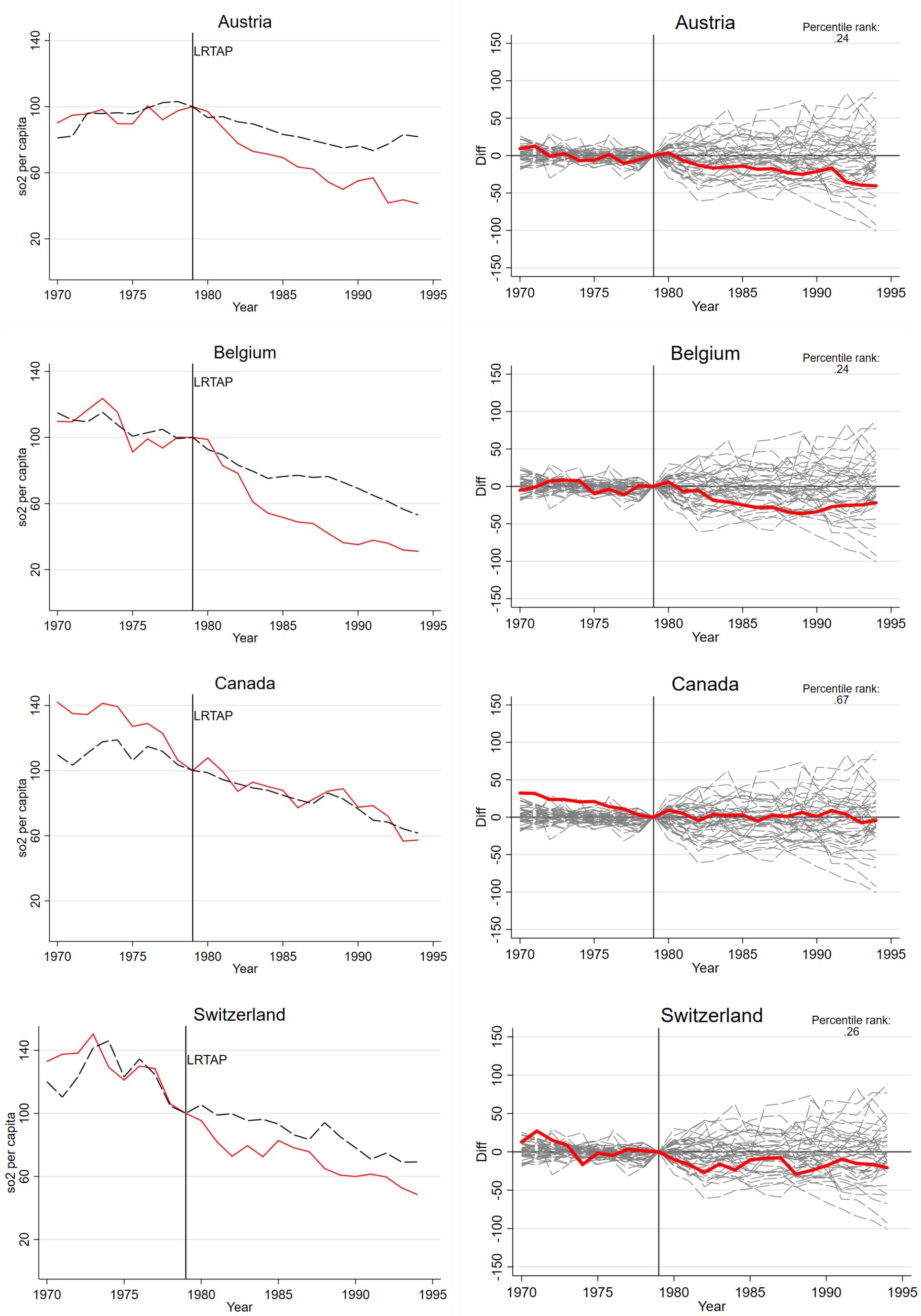


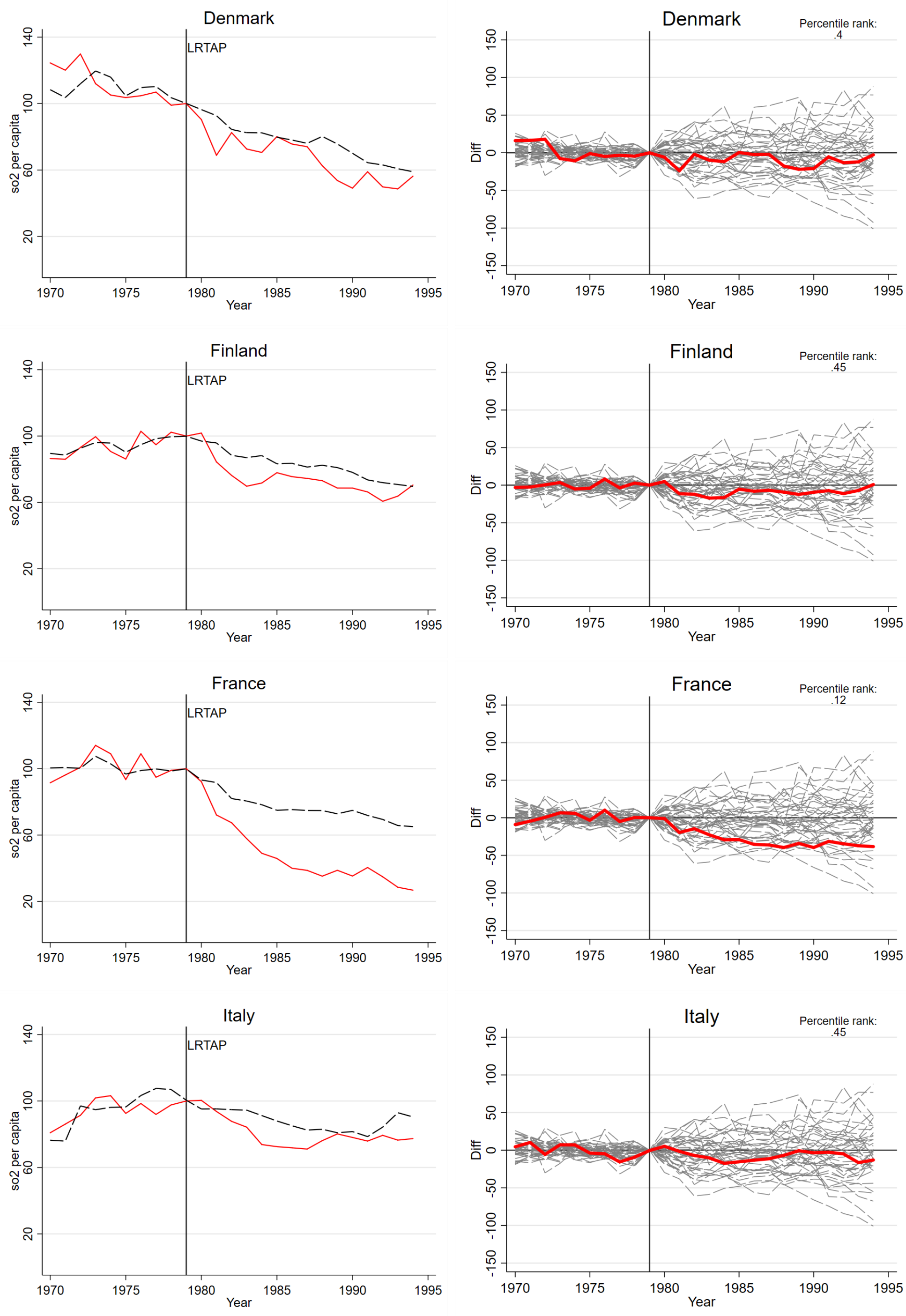
Notes: Figures in row 1 plot the raw mean of log SO_2 emissions, by treatment and control group, for three different samples (*All donors*, *Excluding LRTAP*, *LRTAP only*). Figures in row 2 and 3 plot the coefficients β_{-m} and β_{-m} estimated from equation 11. Shaded area indicate a 95% confidence interval. Figures in row 4 and 5 plot the coefficients β_{-m} and β_{-m} estimated from equation 12. All Figures in column 1 uses the sample *All donors*. All Figures in column 2 uses the sample *Excluding LRTAP*. All Figures in column 3 uses the sample *LRTAP only*. The two last rows include a country-specific linear trend. This means that the estimated treatment effect is measured as deviations from a country-specific linear trend. Standard errors are clustered at the country level in all specifications.

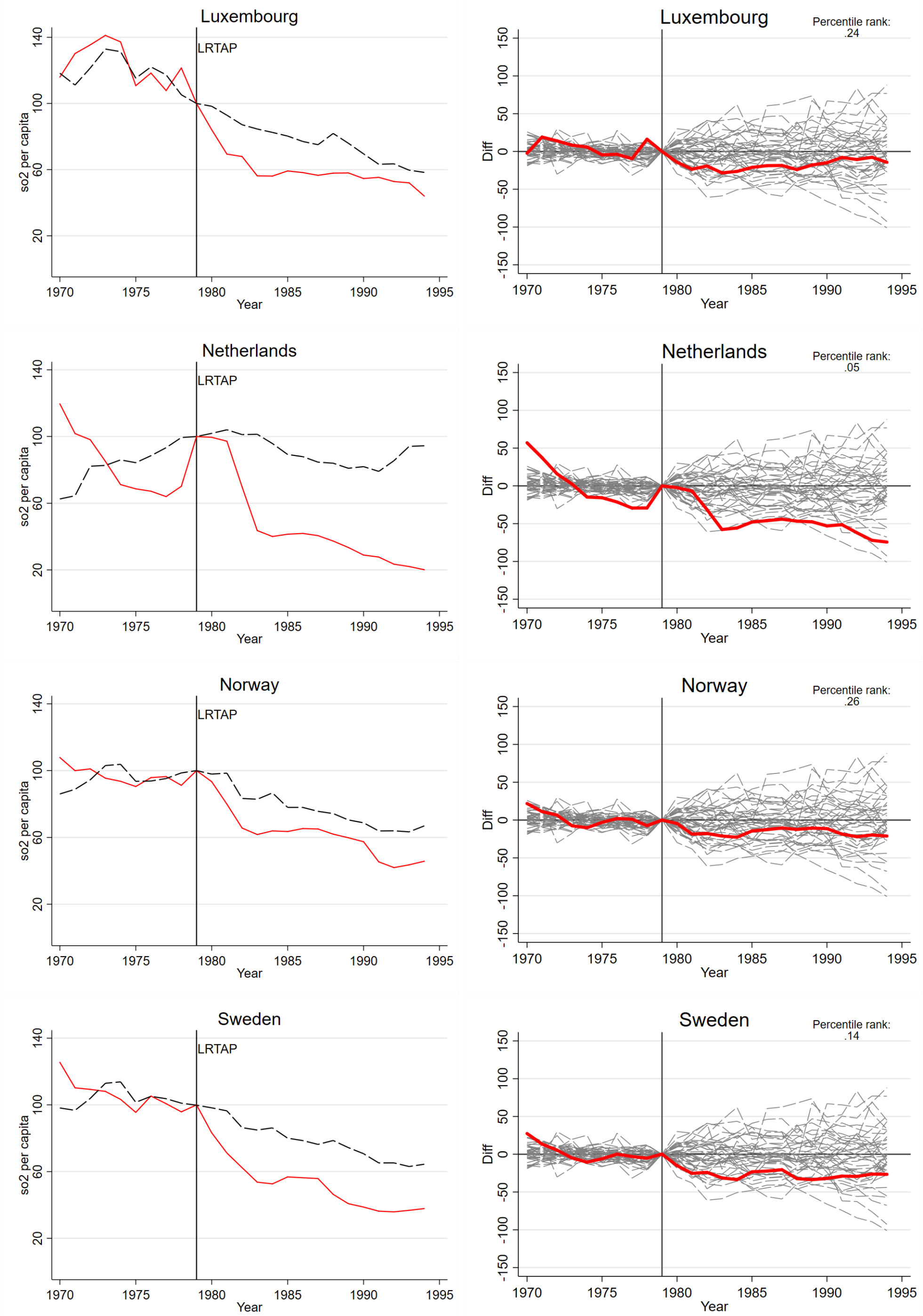
Appendix F Country-level results

F.1 Country estimates and placebo runs: SO_2

Figure F.1: Effects of Helsinki on SO_2 emissions.




Denmark
LRTAP
so2 per capita
20
60
100
140
1970
1975
1980
1985
1990
1995
Year
Denmark
Percentile rank:
.4
Diff
-150
-100
-50
0
50
100
150
Finland
LRTAP
Finland
Percentile rank:
.45
France
LRTAP
France
Percentile rank:
.12
Italy
LRTAP
Italy
Percentile rank:
.45



Notes: Left hand-side figures show the development in emissions for the treated country (red line) and the synthetic control (black, dashed line). Emissions in the year before treatment is normalized to 100. Right hand-side figures show the difference between the treated and synthetic control outcomes $\hat{\alpha}_{1t}$ from equation 3 (thick, red line) and placebo runs (thin, gray lines).

Table F.1: Average country-specific treatment effects. SO_2. 1980-1994

	Mean	p-value
Austria	-19.90	0.24
Belgium	-22.00	0.24
Canada	1.69	0.67
Denmark	-10.11	0.40
Finland	-8.56	0.45
France	-29.48	0.12
Italy	-7.97	0.45
Luxembourg	-17.83	0.24
Netherlands	-46.60	0.05
Norway	-15.88	0.26
Sweden	-26.94	0.14
Switzerland	-16.87	0.26

Table F.2: Country-specific treatment effects. SO_2. 1980-1994.

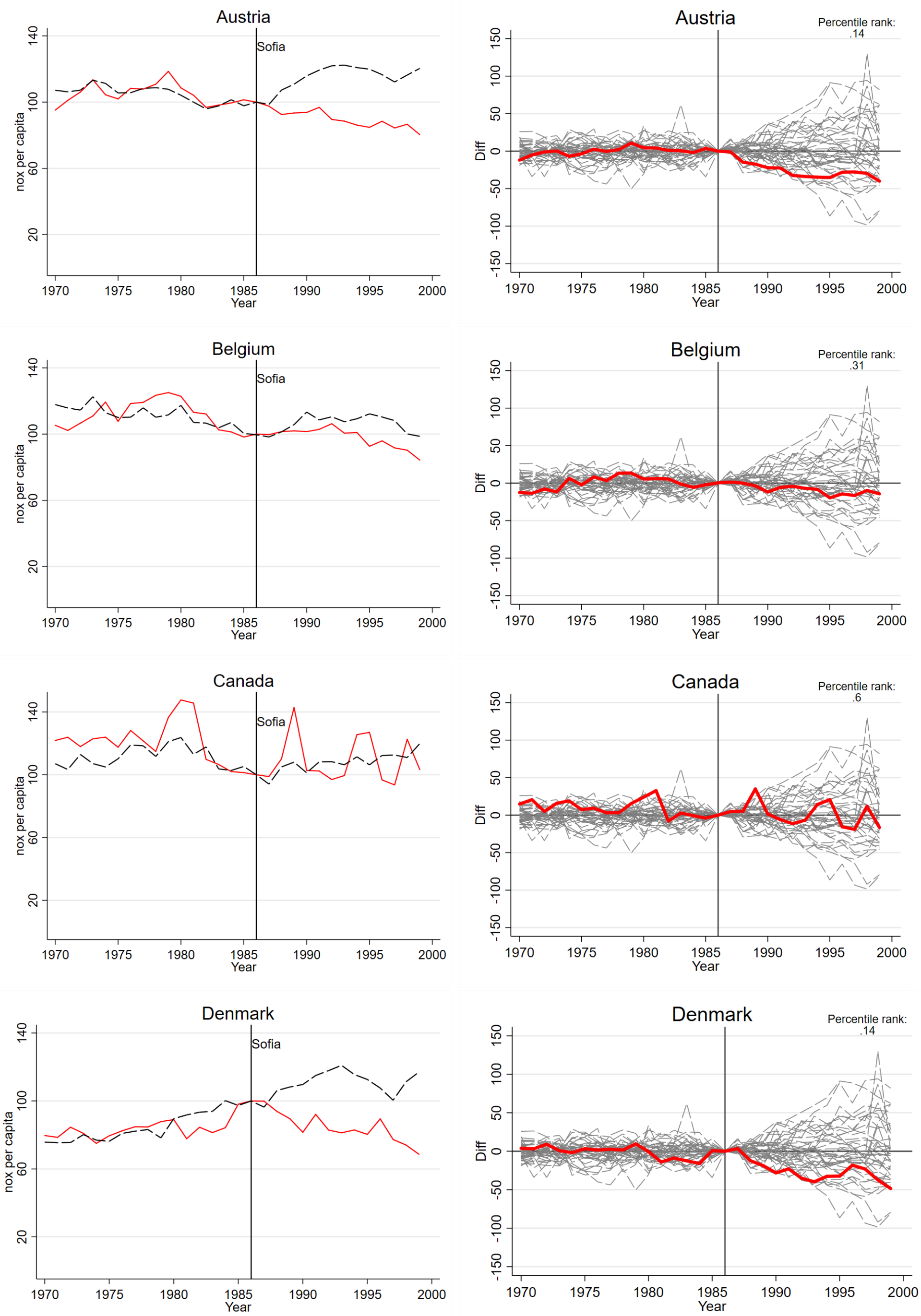
	AUT	BEL	CAN	CHE	DNK	FIN	FRA	ITA	LUX	NLD	NOR	SWE
1980	3.7	6.1	9.1	-10.0	-6.0	4.8	-1.0	5.2	-14.2	-2.4	-4.5	-15.0
1981	-6.7	-6.6	5.0	-16.7	-23.9	-11.5	-19.6	-1.5	-23.5	-6.8	-18.6	-25.3
1982	-12.9	-5.0	-4.5	-26.9	-1.9	-12.1	-14.6	-7.0	-19.2	-31.4	-17.7	-24.0
1983	-16.6	-18.6	3.5	-15.8	-9.8	-17.2	-22.7	-10.2	-28.3	-57.8	-21.1	-31.3
1984	-15.2	-21.1	2.3	-23.7	-11.8	-16.5	-29.3	-17.5	-26.4	-55.7	-22.7	-33.6
1985	-14.0	-24.7	3.0	-10.4	0.3	-5.4	-29.1	-15.5	-21.1	-47.8	-14.4	-23.2
1986	-18.3	-28.3	-4.9	-8.2	-2.4	-8.0	-35.4	-13.3	-18.8	-46.0	-12.7	-22.3
1987	-17.3	-27.9	2.7	-7.7	-2.0	-6.9	-36.1	-11.4	-18.5	-44.1	-10.6	-20.5
1988	-22.8	-34.3	1.1	-29.1	-17.6	-9.3	-39.5	-6.9	-24.0	-46.7	-12.3	-32.3
1989	-25.1	-36.5	6.3	-24.4	-22.0	-12.3	-34.0	-0.7	-17.9	-47.5	-10.7	-33.6
1990	-21.3	-34.0	1.1	-18.3	-21.0	-9.4	-39.6	-3.5	-14.9	-53.0	-11.3	-32.0
1991	-16.5	-27.2	8.9	-9.5	-5.5	-7.3	-31.4	-2.7	-7.9	-51.4	-18.5	-28.9
1992	-35.6	-25.3	3.6	-15.2	-13.2	-11.2	-34.6	-4.9	-10.7	-62.0	-22.0	-29.4
1993	-39.4	-24.7	-7.5	-16.4	-12.1	-7.0	-37.2	-16.5	-7.7	-72.0	-19.7	-26.3
1994	-40.4	-22.0	-4.2	-20.6	-2.7	0.8	-38.3	-13.1	-14.3	-74.3	-21.2	-26.6

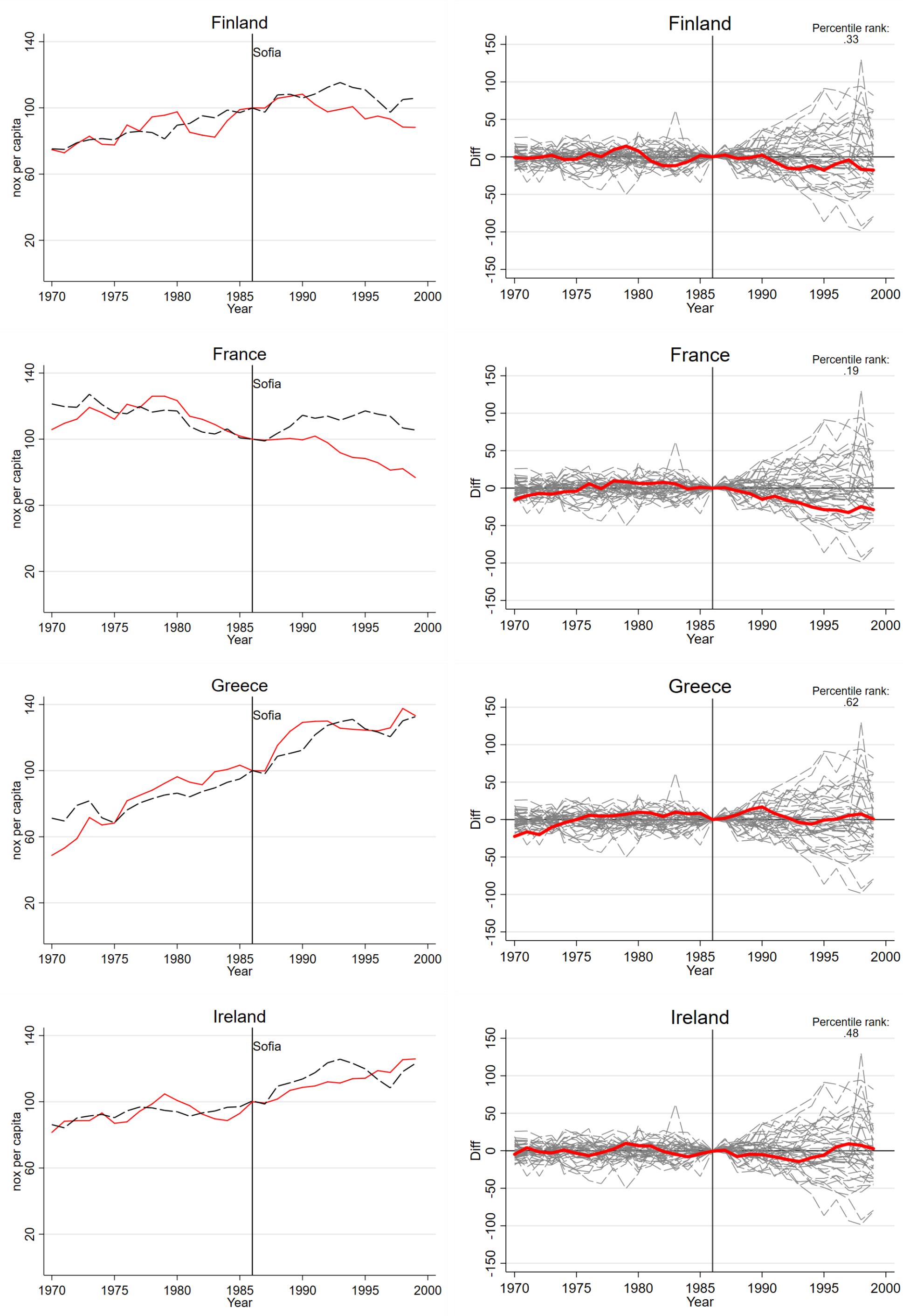
Table F.3: Country-specific p-values. SO_2. 1980-1994.

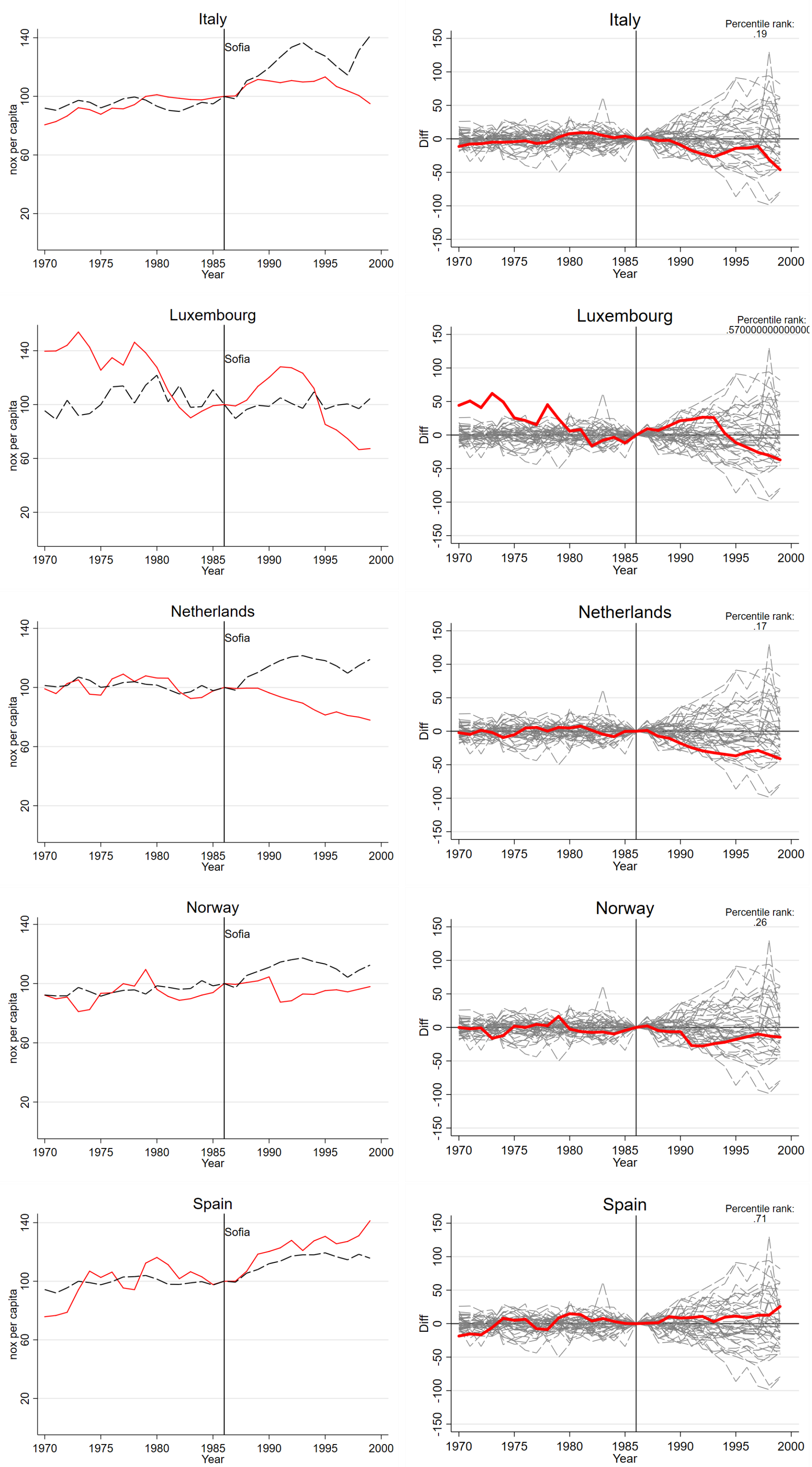
	AUT	BEL	CAN	CHE	DNK	FIN	FRA	ITA	LUX	NLD	NOR	SWE
1980	0.62	0.74	0.74	0.21	0.31	0.71	0.52	0.74	0.12	0.43	0.33	0.12
1981	0.43	0.43	0.62	0.24	0.17	0.36	0.21	0.50	0.17	0.43	0.24	0.17
1982	0.38	0.52	0.52	0.17	0.52	0.38	0.33	0.50	0.29	0.14	0.31	0.24
1983	0.26	0.26	0.57	0.29	0.38	0.26	0.24	0.33	0.19	0.05	0.24	0.14
1984	0.29	0.19	0.64	0.19	0.33	0.26	0.17	0.26	0.17	0.02	0.19	0.12
1985	0.33	0.29	0.67	0.40	0.60	0.48	0.17	0.33	0.31	0.05	0.33	0.31
1986	0.26	0.17	0.45	0.40	0.55	0.40	0.10	0.29	0.26	0.05	0.29	0.26
1987	0.29	0.17	0.67	0.40	0.52	0.43	0.10	0.31	0.29	0.05	0.31	0.26
1988	0.26	0.12	0.67	0.24	0.33	0.50	0.10	0.57	0.26	0.02	0.38	0.17
1989	0.24	0.12	0.64	0.24	0.24	0.38	0.12	0.60	0.31	0.05	0.40	0.12
1990	0.26	0.12	0.62	0.31	0.26	0.43	0.07	0.57	0.38	0.05	0.43	0.17
1991	0.36	0.24	0.67	0.48	0.52	0.50	0.19	0.52	0.50	0.10	0.36	0.21
1992	0.24	0.31	0.60	0.43	0.43	0.43	0.24	0.48	0.43	0.07	0.36	0.29
1993	0.17	0.29	0.48	0.40	0.43	0.48	0.19	0.40	0.48	0.07	0.33	0.29
1994	0.17	0.24	0.50	0.29	0.52	0.60	0.17	0.38	0.33	0.07	0.26	0.24

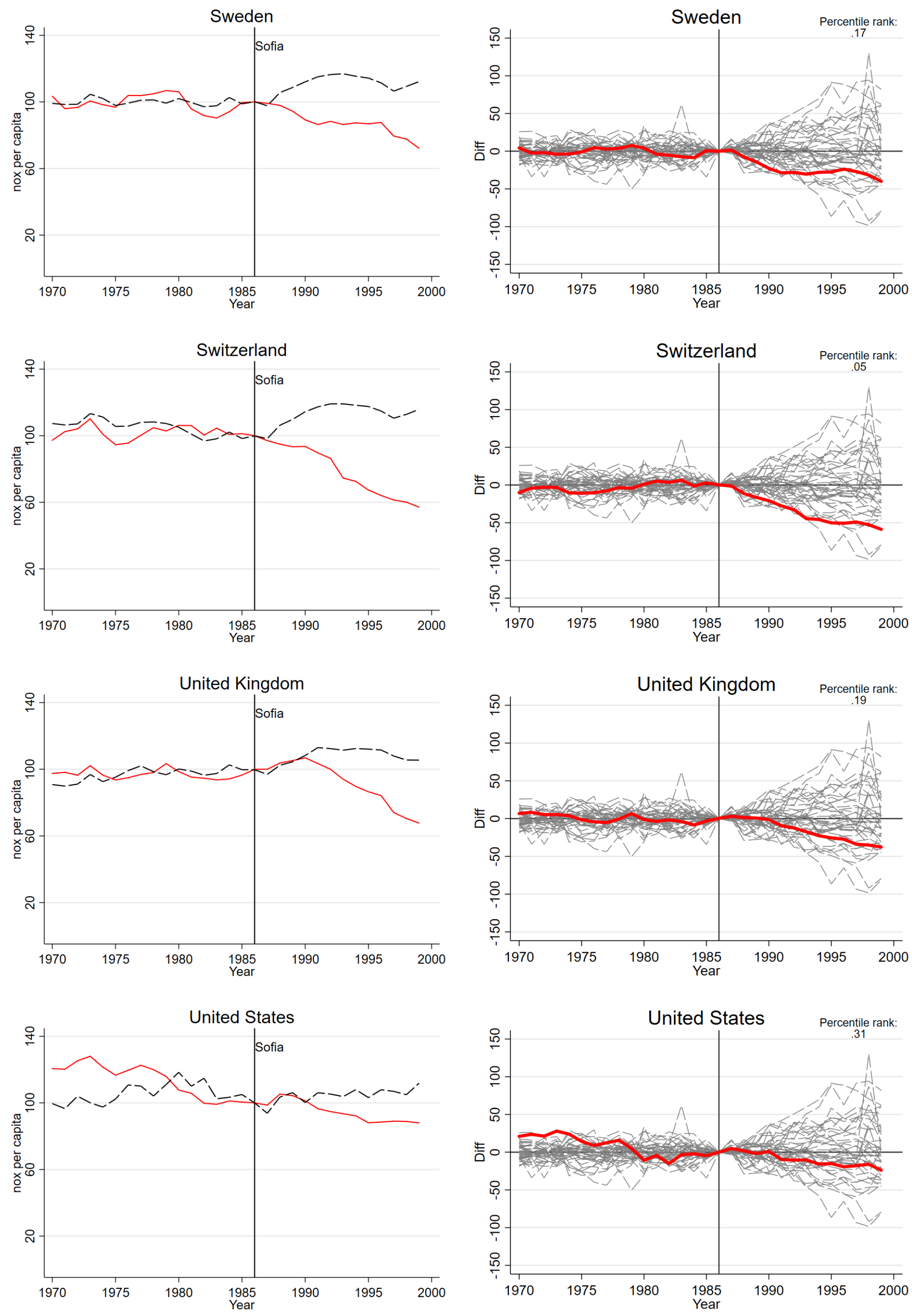
F.2 Country estimates and placebo runs: NO_x

Figure F.0: Effects of Sofia on NO_x emissions.




Finland
Sofia
nox per capita
Year
Diff
Percentile rank:
.33
France
Sofia
Percentile rank:
.19
Greece
Sofia
Percentile rank:
.62
Ireland
Sofia
Percentile rank:
.48
20 60 100 140
-150 -100 -50 0 50 100 150
1970 1975 1980 1985 1990 1995 2000


Italy
Sofia
nox per capita
Year
Italy
Percentile rank:
.19
Diff
Year
Luxembourg
Sofia
nox per capita
Year
Luxembourg
Percentile rank:
.570000000000000
Diff
Year
Netherlands
Sofia
nox per capita
Year
Netherlands
Percentile rank:
.17
Diff
Year
Norway
Sofia
nox per capita
Year
Norway
Percentile rank:
.26
Diff
Year
Spain
Sofia
nox per capita
Year
Spain
Percentile rank:
.71
Diff
Year



Notes: Left hand-side figures show the development in emissions for the treated country (red line) and the synthetic control (black, dashed line). Emissions in the year before treatment is normalized to 100. Right hand-side figures show the difference between the treated and synthetic control outcomes $\hat{\alpha}_{1t}$ from equation 3 (thick, red line) and placebo runs (thin, gray lines).

Table F.4: Average country-specific treatment effects. NO_x. 1987-1999

	Mean	p-value
Austria	-26.12	0.14
Belgium	-8.75	0.31
Canada	1.35	0.60
Denmark	-26.54	0.14
Finland	-8.66	0.33
France	-18.54	0.19
Greece	4.06	0.62
Ireland	-3.15	0.48
Italy	-16.58	0.19
Luxembourg	0.53	0.57
Netherlands	-25.22	0.17
Norway	-14.28	0.26
Spain	9.53	0.71
Sweden	-23.71	0.17
Switzerland	-35.53	0.05
United Kingdom	-16.76	0.19
United States	-10.20	0.31

Table F.5: Country-specific treatment effects. NO_x. 1987.

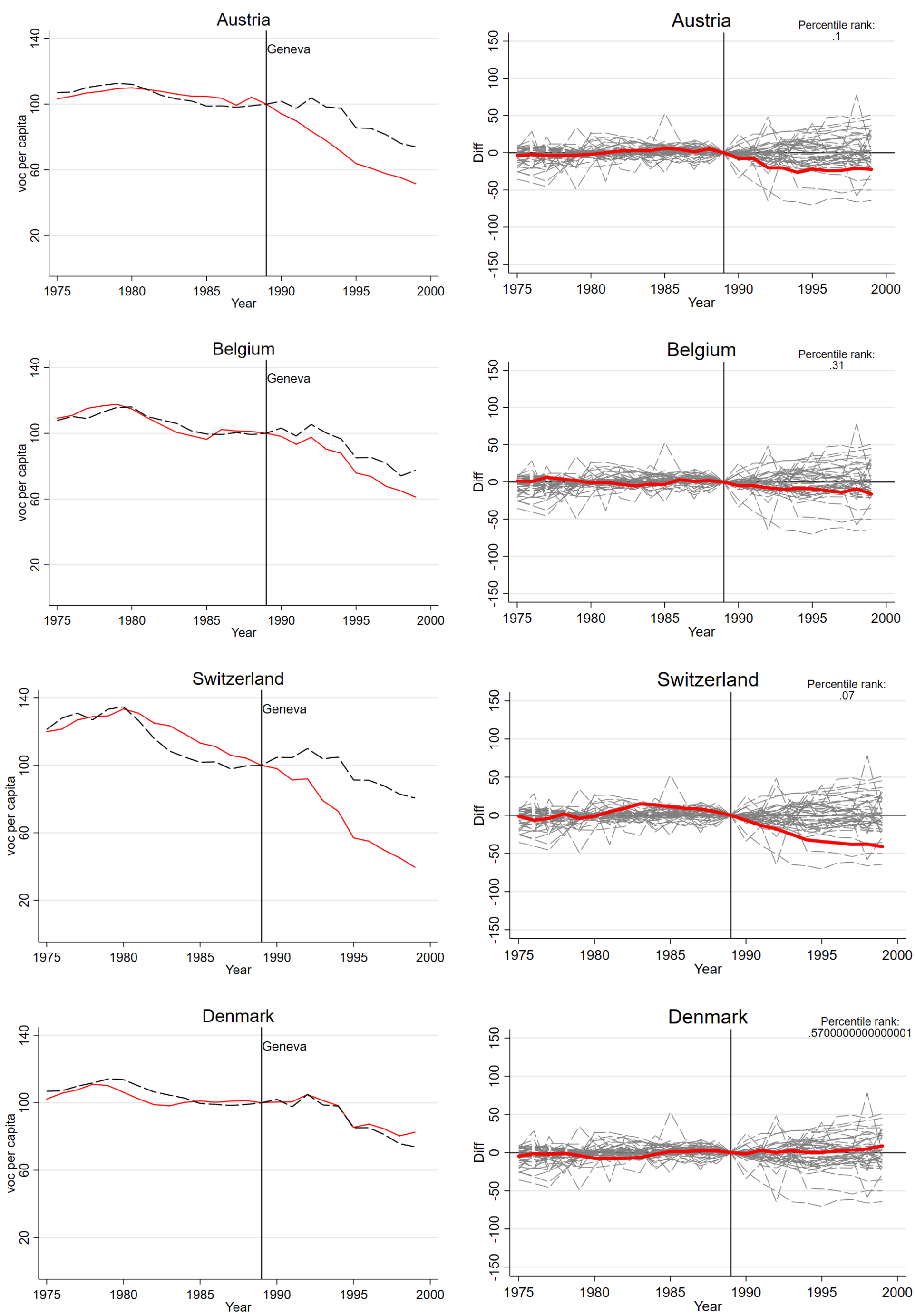
	AUT	BEL	CAN	CHE	DNK	ESP	FIN	FRA	GBR	GRC	IRL	ITA	LUX	NLD	NOR	SWE	USA
1987	-1.1	1.4	4.8	-1.2	3.5	0.7	2.5	0.5	3.1	1.8	0.7	2.0	9.3	1.1	2.3	1.5	4.9
1988	-14.8	0.1	5.1	-11.4	-12.3	1.3	-2.1	-3.6	1.3	6.5	-7.7	-2.6	6.8	-7.4	-4.8	-7.6	1.8
1989	-17.4	-3.8	34.9	-16.4	-18.7	10.5	-1.2	-7.2	0.7	13.3	-4.6	-2.3	14.0	-10.7	-6.4	-14.2	-1.8
1990	-22.1	-11.8	1.5	-20.8	-28.1	8.3	2.3	-14.9	-1.4	16.8	-5.1	-9.0	21.5	-18.1	-6.4	-22.8	1.1
1991	-22.5	-5.8	-5.8	-27.6	-22.9	9.1	-6.4	-10.8	-9.5	8.1	-8.0	-17.5	23.0	-24.5	-27.1	-28.7	-9.6
1992	-32.4	-4.2	-11.4	-32.8	-35.0	10.7	-14.8	-16.1	-12.4	2.6	-11.5	-22.7	26.5	-29.3	-27.7	-28.1	-10.5
1993	-33.8	-6.9	-6.9	-44.6	-39.7	2.9	-16.2	-19.6	-17.4	-3.9	-14.4	-26.9	26.1	-32.1	-24.3	-30.5	-10.3
1994	-34.7	-8.3	14.1	-45.7	-32.6	9.5	-11.6	-25.1	-22.7	-6.0	-9.3	-20.9	2.7	-34.5	-22.1	-28.0	-15.7
1995	-35.1	-19.6	20.7	-50.2	-32.3	11.3	-17.6	-28.8	-25.6	-0.7	-5.6	-14.2	-11.3	-36.8	-18.0	-27.5	-15.1
1996	-28.2	-14.4	-15.6	-50.7	-18.1	8.8	-9.2	-29.4	-27.3	0.7	5.3	-13.8	-18.4	-31.1	-14.0	-23.8	-19.4
1997	-27.8	-16.5	-19.0	-49.1	-23.1	12.5	-4.1	-32.6	-33.9	5.4	9.3	-10.8	-25.9	-28.7	-9.9	-27.0	-17.9
1998	-29.7	-9.8	11.6	-52.7	-37.2	12.7	-16.6	-24.6	-35.1	7.5	7.2	-30.7	-30.4	-34.9	-12.8	-31.6	-16.0
1999	-40.0	-14.3	-16.5	-58.8	-48.4	25.6	-17.6	-28.7	-37.7	0.7	2.8	-46.2	-37.1	-41.0	-14.5	-40.0	-23.9

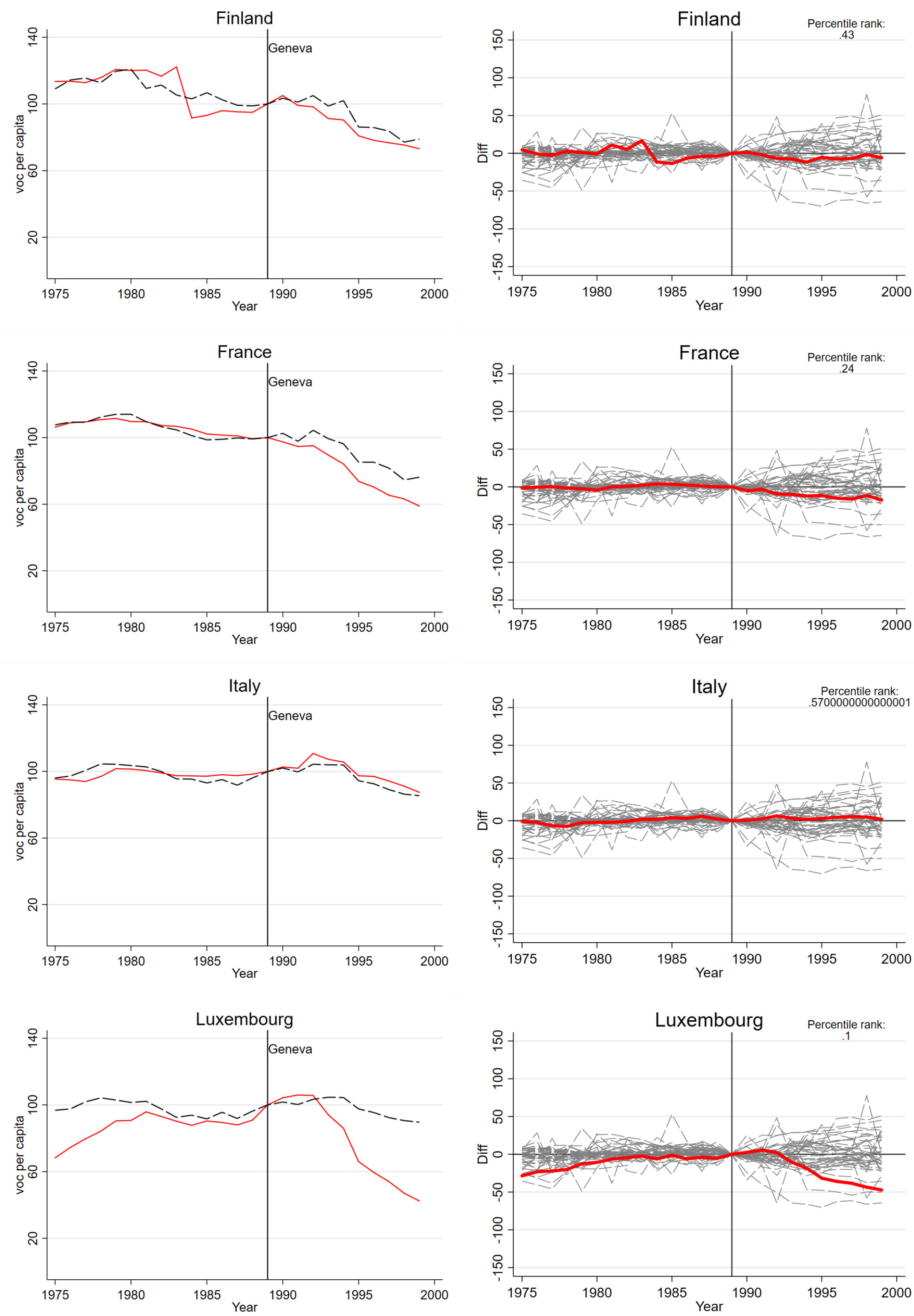
Table F.6: Country-specific p-values. NO_x. 1987.

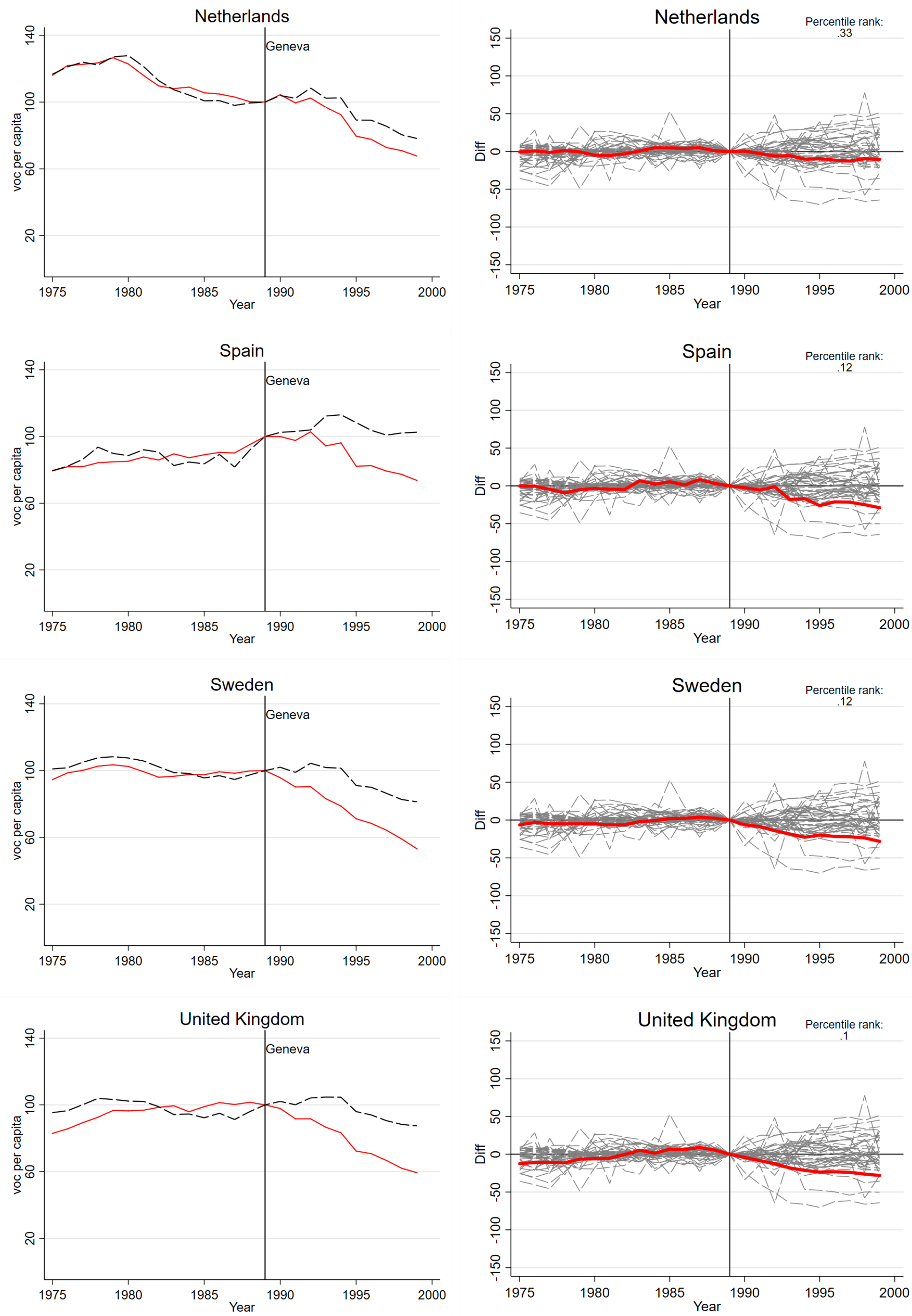
	AUT	BEL	CAN	CHE	DNK	ESP	FIN	FRA	GBR	GRC	IRL	ITA	LUX	NLD	NOR	SWE	USA
1987	0.19	0.64	0.81	0.17	0.76	0.50	0.71	0.43	0.71	0.64	0.50	0.67	0.93	0.62	0.67	0.64	0.81
1988	0.14	0.45	0.64	0.14	0.14	0.50	0.33	0.24	0.50	0.69	0.19	0.31	0.69	0.19	0.24	0.19	0.50
1989	0.14	0.43	1.00	0.17	0.14	0.81	0.45	0.31	0.57	0.93	0.40	0.45	0.93	0.21	0.36	0.17	0.45
1990	0.12	0.17	0.55	0.12	0.05	0.69	0.55	0.17	0.45	0.86	0.36	0.31	0.86	0.12	0.33	0.12	0.52
1991	0.10	0.36	0.36	0.07	0.10	0.71	0.33	0.29	0.29	0.69	0.33	0.14	0.86	0.07	0.07	0.07	0.29
1992	0.05	0.40	0.33	0.05	0.05	0.71	0.24	0.24	0.29	0.64	0.33	0.14	0.86	0.07	0.07	0.07	0.33
1993	0.10	0.38	0.38	0.07	0.10	0.52	0.26	0.19	0.26	0.40	0.26	0.10	0.83	0.10	0.14	0.10	0.33
1994	0.17	0.45	0.69	0.05	0.17	0.64	0.43	0.17	0.19	0.48	0.45	0.19	0.60	0.17	0.19	0.17	0.36
1995	0.17	0.24	0.74	0.05	0.17	0.67	0.26	0.21	0.21	0.55	0.52	0.33	0.43	0.14	0.26	0.21	0.29
1996	0.19	0.31	0.26	0.10	0.26	0.67	0.40	0.19	0.19	0.55	0.62	0.36	0.26	0.19	0.33	0.19	0.26
1997	0.24	0.31	0.26	0.14	0.24	0.69	0.57	0.21	0.21	0.64	0.69	0.40	0.24	0.24	0.43	0.24	0.26
1998	0.24	0.43	0.60	0.10	0.17	0.62	0.31	0.29	0.17	0.60	0.60	0.24	0.24	0.17	0.38	0.24	0.31
1999	0.17	0.38	0.38	0.07	0.07	0.79	0.38	0.29	0.17	0.60	0.60	0.07	0.17	0.14	0.38	0.17	0.31

F.3 Country estimates and placebo runs: VOC

Figure F.-2: Effects of Geneva on VOC emissions.




Finland
Geneva
voc per capita
Year
Finland
Percentile rank:
.43
Diff
Year
France
Geneva
voc per capita
Year
France
Percentile rank:
.24
Diff
Year
Italy
Geneva
voc per capita
Year
Italy
Percentile rank:
.5700000000000001
Diff
Year
Luxembourg
Geneva
voc per capita
Year
Luxembourg
Percentile rank:
.1
Diff
Year



Notes: Left hand-side figures show the development in emissions for the treated country (red line) and the synthetic control (black, dashed line). Emissions in the year before treatment is normalized to 100. Right hand-side figures show the difference between the treated and synthetic control outcomes $\hat{\alpha}_{1t}$ from equation 3 (thick, red line) and placebo runs (thin, gray lines).

Table F.7: Average country-specific treatment effects. VOC. 1990-1999.

	Mean	p-value
Austria	-19.57	0.10
Belgium	-9.71	0.31
Denmark	2.35	0.57
Finland	-5.37	0.43
France	-11.08	0.24
Italy	3.38	0.57
Luxembourg	-21.51	0.10
Netherlands	-7.78	0.33
Spain	-16.64	0.12
Sweden	-18.57	0.12
Switzerland	-28.27	0.07
United Kingdom	-18.98	0.10

Table F.8: Country-specific treatment effects. VOC. 1990-1999.

	AUT	BEL	CHE	DNK	ESP	FIN	FRA	GBR	ITA	LUX	NLD	SWE
1990	-7.7	-5.1	-6.8	-1.6	-2.4	1.6	-5.1	-4.2	0.6	2.6	0.5	-6.2
1991	-7.6	-5.0	-13.3	3.0	-5.4	-2.0	-3.1	-8.5	2.2	5.7	-2.7	-8.7
1992	-20.3	-7.8	-17.9	-0.0	-1.2	-6.7	-9.2	-12.5	6.5	2.3	-6.0	-14.0
1993	-20.5	-9.9	-24.9	2.7	-17.9	-7.5	-9.9	-18.1	3.3	-10.5	-5.4	-18.6
1994	-26.4	-8.8	-32.1	0.2	-16.8	-11.6	-12.1	-21.3	1.8	-18.4	-10.1	-22.7
1995	-21.8	-9.2	-34.4	0.3	-26.1	-5.3	-11.5	-23.7	3.0	-31.6	-9.7	-19.9
1996	-24.2	-11.7	-36.1	2.1	-21.2	-7.6	-14.9	-23.2	4.4	-35.9	-11.5	-21.6
1997	-23.8	-14.1	-38.2	3.4	-21.5	-7.0	-16.3	-23.9	5.2	-38.4	-12.8	-22.1
1998	-20.9	-9.2	-37.8	4.7	-24.9	-1.7	-11.4	-26.3	4.8	-43.6	-9.5	-23.6
1999	-22.4	-16.3	-41.2	8.7	-28.9	-5.8	-17.3	-28.1	2.0	-47.2	-10.5	-28.2

Table F.9: Country-specific p-values. VOC. 1990-1999.

	AUT	BEL	CHE	DNK	ESP	FIN	FRA	GBR	ITA	LUX	NLD	SWE
1990	0.12	0.21	0.12	0.50	0.43	0.76	0.21	0.31	0.69	0.81	0.69	0.17
1991	0.21	0.38	0.10	0.74	0.38	0.48	0.45	0.19	0.69	0.81	0.45	0.19
1992	0.10	0.29	0.12	0.57	0.55	0.36	0.26	0.19	0.83	0.69	0.45	0.17
1993	0.07	0.24	0.05	0.64	0.10	0.31	0.24	0.10	0.67	0.24	0.38	0.10
1994	0.10	0.38	0.07	0.48	0.19	0.36	0.33	0.12	0.48	0.19	0.38	0.12
1995	0.14	0.33	0.07	0.52	0.07	0.45	0.33	0.12	0.57	0.07	0.33	0.14
1996	0.10	0.26	0.07	0.57	0.10	0.31	0.19	0.10	0.57	0.07	0.26	0.10
1997	0.10	0.14	0.07	0.57	0.10	0.36	0.14	0.10	0.62	0.07	0.17	0.10
1998	0.14	0.38	0.10	0.67	0.12	0.48	0.31	0.12	0.67	0.10	0.38	0.12
1999	0.14	0.24	0.07	0.71	0.10	0.43	0.24	0.12	0.57	0.07	0.33	0.12

Appendix G Robustness checks: country weights

This appendix lists the weights used for constructing the synthetic controls in the robustness checks presented in Table 10 and Figure 9.

G.1 Weights used for constructing synthetic controls. SO_2. 1980

Table G.1: Robustness R1

	AUT	BEL	CAN	CHE	DNK	FIN	FRA	ITA	LUX	NLD	NOR	SWE
Australia						21.6						
Bahrain										2.6		
Chile			28.1			20.3						5.3
ElSalvador												1.7
Greece								9.2				
Iceland	27.1				9.3	31.3	17.5	10.4	1.4	34.1	78.4	42.7
Japan	63.5	22.8					23.4	33.8		41.0		
KoreaRep	7.9											
NewZealand		11.1	32.4	85.6	19.4	3.7	3.3		43.2		8.6	19.6
SaudiArabia	1.4							1.8		22.3		
Singapore		1.9					5.8	12.6				
UnitedArabEmirates				13.4								
UnitedKingdom		4.8				23.2	48.4	32.1				4.4
UnitedStates		59.4	39.5	1.0	71.4		1.6		55.5		12.9	26.3

Table G.2: Robustness R2

	AUT	BEL	CAN	CHE	DNK	FIN	FRA	ITA	LUX	NLD	NOR	SWE
Algeria	0.8	0.2				0.3	0.5	0.9		1.0		0.1
Argentina	1.0	0.4				0.4	0.8	1.3		0.6		0.2
Australia	0.3	0.6				20.1	1.2	1.3		0.2		0.2
Bahrain	8.6	0.4				0.6	1.2	1.9		2.0		0.3
Brazil	0.8	0.3				0.4	0.6	1.0		0.8		0.1
Chile	0.2		26.9			4.6	0.3	0.4		0.2		
Colombia	1.0	0.3				0.3	0.7	1.1		0.6		0.2
CostaRica	0.8	0.2				0.3	0.6	1.0		0.9		0.1
CotedIvoire	1.0	0.4				0.3	0.7	1.1		0.5		0.2
Cuba	0.7	0.3				0.4	0.6	1.0		0.6		0.1
Cyprus	0.9	0.4				0.6	0.9	1.3		0.8		0.2
Ecuador	0.7	0.1				0.3	0.4	0.8		1.2		0.1
ElSalvador	1.4	0.5				0.3	0.7	1.5		0.4		0.3
Greece	0.8	0.4				0.9	1.0	1.5		0.8		0.3
Honduras	0.8	0.2				0.3	0.6	1.0		0.7		0.1
HongKongSAR	0.7	0.4				0.7	0.9	1.3		0.6		0.2
Iceland	50.8	10.3			9.8	13.7	35.9	33.7	2.1	61.7	78.8	35.7
IranRep	0.8	0.2				0.3	0.6	0.9		0.7		0.1
Ireland	0.7	0.5				1.2	1.1	1.6		0.6		0.3
Israel	0.7	0.5				0.9	0.9	1.3		0.6		0.2
Japan	1.8	0.7				0.9	1.8	2.7		0.9		0.5
KoreaRep	0.7	0.2				0.4	0.5	0.9		0.9		0.1
Malaysia	0.7	0.2				0.4	0.5	0.9		0.8		0.1
Malta	0.7	0.3				0.6	0.7	1.1		0.7		0.2
Mexico	0.9	0.3				0.5	0.8	1.2		0.7		0.2
NewZealand	9.8	16.7	31.7	85.6	19.3	0.6	16.7	6.2	42.9	0.6	8.5	18.7
Nicaragua	1.1	0.5				0.3	0.8	1.3		0.4		0.2
Panama	0.9	0.4				0.4	0.7	1.2		0.6		0.2
Paraguay	0.8	0.2				0.3	0.6	0.9		0.7		0.1
Peru	0.6	0.2				0.5	0.5	0.9		0.6		0.1
Portugal	0.9	0.3				0.5	0.7	1.2		1.0		0.2
SaudiArabia	0.8	0.3				2.2	0.8	1.5		10.5		0.2
Singapore	0.7	0.3				0.7	0.7	1.2		0.8		0.2
SouthAfrica	0.6	0.4				0.6	0.7	1.1		0.5		0.2
Spain	0.8	0.5				0.9	1.1	1.6		0.7		0.3
SyrianArabRepublic	0.7	0.2				0.3	0.5	0.8		0.9		0.1
Tunisia	0.7	0.2				0.3	0.5	0.9		0.8		0.1
Turkey	0.9	0.3				0.4	0.7	1.1		0.7		0.2
UnitedArabEmirates				13.7								
UnitedKingdom	1.1	1.0				1.3	2.1	3.0		0.6		0.7
UnitedStates	0.3	59.6	41.4	0.7	71.0	40.4	17.7	13.4	55.0	0.2	12.7	37.9
Venezuela	0.8	0.2				0.4	0.6	1.0		0.9		0.1

Table G.3: Robustness R3

	AUT	BEL	CAN	CHE	DNK	FIN	FRA	ITA	LUX	NLD	NOR	SWE
Algeria	1.9	0.4		0.2	0.2	0.6	1.1	1.5		2.3	1.3	0.7
Argentina	2.6	1.1		0.5	0.5	1.2	2.1	2.4		1.9	2.3	1.6
Australia	2.5	2.4		0.9	1.2	4.0	3.2	3.5		2.8	3.6	2.9
Bahrain	2.2	0.6		0.3	0.3	0.8	1.5	1.8		2.1	1.6	1.0
Brazil	2.1	0.6		0.3	0.2	0.8	1.4	1.8		2.2	1.5	0.9
Chile	2.6	31.2	56.9	6.8	40.0	37.4	11.3	8.5	33.4	4.5	11.7	20.8
Colombia	2.5	0.9		0.4	0.4	1.0	1.9	2.2		1.8	2.1	1.4
CostaRica	2.1	0.5		0.2	0.2	0.7	1.3	1.7		2.2	1.4	0.8
CotedIvoire	2.7	1.1		0.4	0.5	1.1	2.2	2.5		1.8	2.3	1.7
Cuba	2.2	0.7		0.3	0.3	1.2	1.6	2.0		2.3	1.8	1.1
Cyprus	2.2	0.8		0.3	0.3	1.3	1.6	2.0		2.3	1.8	1.1
Ecuador	1.8	0.3		0.2	0.1	0.6	1.0	1.4		2.6	1.1	0.5
ElSalvador	5.9	1.8		31.5	0.8	1.2	5.5	3.7		1.5	3.8	3.6
Greece	2.1	0.8		0.3	0.4	1.5	1.6	2.0		2.5	1.8	1.1
Honduras	2.2	0.6		0.3	0.3	0.8	1.5	1.8		2.1	1.6	0.9
HongKongSAR	2.3	1.0		0.4	0.5	1.7	1.9	2.3		2.3	2.1	1.5
Iceland	2.3	0.9		0.4	0.4	1.5	1.8	2.2		2.3	2.0	1.4
IranRep	2.1	0.5		0.3	0.2	0.8	1.4	1.8		2.2	1.5	0.9
Ireland	2.3	1.2		0.5	0.5	2.1	2.0	2.4		2.5	2.3	1.6
Israel	2.4	1.2		0.5	0.6	2.0	2.1	2.5		2.3	2.3	1.7
Japan	2.5	1.1		0.5	0.5	1.4	2.1	2.4		2.0	2.3	1.6
KoreaRep	1.9	0.4		0.2	0.2	0.8	1.2	1.6		2.6	1.3	0.7
Malaysia	2.0	0.5		0.2	0.2	0.8	1.3	1.7		2.3	1.4	0.8
Malta	2.1	0.7		0.3	0.3	1.3	1.6	2.0		2.4	1.8	1.1
Mexico	2.2	0.7		0.3	0.3	1.1	1.6	2.0		2.1	1.8	1.1
NewZealand	6.0	31.7	43.1	47.2	42.4	10.6	15.2	6.8	66.6	1.6	9.9	24.3
Nicaragua	3.3	1.5		0.3	0.7	1.1	3.0	3.0		1.6	3.0	2.4
Panama	2.5	1.0		0.4	0.4	1.1	2.0	2.3		1.9	2.1	1.5
Paraguay	2.2	0.6		0.3	0.3	0.8	1.5	1.8		2.1	1.6	1.0
Peru	2.0	0.6		0.3	0.3	1.3	1.4	1.8		2.7	1.6	1.0
Portugal	2.0	0.5		0.2	0.2	1.0	1.3	1.7		2.4	1.5	0.9
SaudiArabia	1.6	0.4		0.1	0.2	1.1	1.0	1.5		9.5	1.2	0.6
Singapore	2.0	0.7		0.3	0.3	1.4	1.5	1.9		2.6	1.7	1.0
SouthAfrica	2.6	1.6		0.7	0.8	2.1	2.5	2.9		2.1	2.8	2.2
Spain	2.3	1.1		0.4	0.5	1.7	2.0	2.3		2.3	2.2	1.5
SyrianArabRepublic	1.9	0.4			0.2	0.7	1.1	1.6		2.4	1.3	0.7
Tunisia	2.0	0.5		0.2	0.2	0.8	1.3	1.7		2.3	1.4	0.8
Turkey	2.3	0.8		0.3	0.3	1.0	1.7	2.0		2.0	1.8	1.2
UnitedArabEmirates				0.1								
UnitedKingdom	2.7	1.7		0.7	0.8	2.2	2.6	2.9		2.1	3.0	2.3
UnitedStates	2.8	4.5		1.7	2.6	4.4	5.0	4.5		2.4	5.0	5.3
Venezuela	2.0	0.5		0.2	0.2	0.9	1.3	1.7		2.4	1.4	0.8

Table G.4: Robustness R4

	AUT	BEL	CAN	CHE	DNK	FIN	FRA	ITA	LUX	NLD	NOR	SWE
Algeria	0.6	0.3				0.7	0.5	0.8		0.7		0.2
Argentina	1.1	0.7				1.3	0.9	1.4		0.5		0.4
Australia	2.0	1.2				2.2	1.7	2.3		1.4		0.8
Bahrain	1.7	1.0				1.9	1.4	2.0		1.3		0.7
Brazil	0.7	0.4				0.9	0.6	1.0		0.6		0.3
Chile	0.7	0.5				0.9	0.6	1.0		0.6		0.3
Colombia	0.9	0.6				1.0	0.8	1.1		0.4		0.3
CostaRica	0.7	0.4				0.8	0.6	0.9		0.6		0.2
CotedIvoire	0.9	0.6				1.0	0.8	1.1		0.4		0.3
Cuba	0.7	0.4				0.8	0.6	0.9		0.6		0.2
Cyprus	1.2	0.7				1.4	1.0	1.5		1.0		0.4
Ecuador	0.5	0.3				0.6	0.4	0.7		0.8		0.1
ElSalvador	1.4	1.0				1.7	1.0	2.0		0.3		0.5
Greece	1.4	0.8				1.6	1.1	1.7		1.4		0.5
Honduras	0.6	0.4				0.8	0.5	0.8		0.5		0.2
HongKongSAR	1.2	0.7				1.4	1.0	1.5		0.9		0.4
Iceland	48.4	40.9	23.2		61.5	42.5	43.4	38.7	52.2	58.9	92.5	54.8
IranRep	0.6	0.4				0.8	0.5	0.8		0.6		0.2
Ireland	1.7	0.9				1.9	1.4	2.0		1.3		0.6
Israel	1.3	0.8				1.5	1.1	1.6		0.9		0.5
Japan	2.5	1.6				2.7	2.2	2.8		1.2		1.1
KoreaRep	0.6	0.3				0.7	0.5	0.8		0.8		0.2
Malaysia	0.6	0.3				0.7	0.5	0.8		0.6		0.2
Malta	0.9	0.5				1.0	0.7	1.1		0.9		0.3
Mexico	1.0	0.6				1.1	0.8	1.2		0.7		0.3
NewZealand	8.4	32.2	76.8	77.6	38.5	7.9	19.9	7.8	47.8	0.7	7.5	28.6
Nicaragua	1.0	0.8				1.3	0.9	1.4		0.3		0.4
Panama	0.9	0.6				1.1	0.8	1.2		0.5		0.3
Paraguay	0.6	0.4				0.7	0.5	0.8		0.5		0.2
Peru	0.6	0.3				0.7	0.5	0.8		0.7		0.2
Portugal	1.0	0.6				1.2	0.8	1.3		1.1		0.3
SaudiArabia	1.1	0.6				1.2	0.8	1.3		11.0		0.3
Singapore	1.0	0.5				1.2	0.8	1.2		1.1		0.3
SouthAfrica	1.1	0.7				1.2	0.9	1.3		0.5		0.4
Spain	1.6	0.9				1.8	1.3	1.9		1.1		0.6
SyrianArabRepublic	0.5	0.3				0.6	0.4	0.7		0.6		0.2
Tunisia	0.6	0.3				0.7	0.5	0.8		0.6		0.2
Turkey	0.8	0.5				1.0	0.7	1.1		0.5		0.3
UnitedArabEmirates				22.4								
UnitedKingdom	2.8	1.9				3.0	2.5	3.1		1.1		1.3
UnitedStates	3.4	2.7				3.6	3.3	3.7		1.1		1.9
Venezuela	0.8	0.4				0.9	0.6	1.0		0.9		0.3

Table G.5: Robustness R5

	AUT	BEL	CAN	CHE	DNK	FIN	FRA	ITA	LUX	NLD	NOR	SWE
Algeria	1.9			0.3		1.9	1.6	1.3		2.4	1.3	1.3
Argentina	3.3			1.4		2.4	2.7	0.3		0.2	1.7	2.0
Australia	2.2			0.5		2.2	2.0	1.1		0.5	1.5	1.6
Bahrain	4.0			0.6		2.1	2.5	20.8		55.7	1.4	1.5
Brazil	1.1			0.2		2.4	0.9	0.4		0.2	1.9	2.0
Chile	1.4			0.3		2.2	1.3	0.6		0.3	1.8	1.8
Colombia	1.6			0.5		2.6	1.6	0.6		0.2	2.1	2.3
CostaRica	1.2			0.2		2.3	0.9	0.5		0.2	1.8	1.8
CotedIvoire	1.2			0.3		3.1	1.1	0.3		0.1	3.2	3.2
Cuba	1.3			0.3		2.3	1.1	0.6		0.2	1.8	1.8
Cyprus	2.7			0.4		2.1	2.1	1.7		2.8	1.4	1.5
Ecuador	1.1			0.1		1.9	0.8	0.6		0.3	1.3	1.3
ElSalvador	0.8			4.9		7.9	0.3	0.1		0.1	14.1	26.2
Greece	1.8			0.3		2.0	1.6	1.1		0.7	1.4	1.4
Honduras	0.9			0.2		2.8	0.7	0.3		0.1	2.2	2.8
HongKongSAR	3.7			0.6		2.1	2.4	2.8		8.5	1.4	1.6
Iceland	1.0			0.2		2.6	0.8	0.3		0.2	2.2	2.5
IranRep	2.3			0.4		2.0	1.9	1.5		1.4	1.4	1.4
Ireland	1.7			0.3		2.2	1.6	0.8		0.3	1.6	1.7
Israel	3.4			0.6		2.2	2.4	1.9		1.9	1.5	1.6
Japan	2.8			0.9		2.3	2.5	0.8		0.3	1.6	1.9
KoreaRep	1.6			0.2		1.9	1.3	1.1		1.1	1.3	1.2
Malaysia	1.5			0.2		2.1	1.2	0.7		0.4	1.5	1.5
Malta	2.3			0.3		2.0	1.9	1.5		5.9	1.4	1.4
Mexico	2.0			0.4		2.2	1.9	1.0		0.4	1.6	1.7
NewZealand	5.7	6.7	74.2	70.9	17.0	3.8	15.5		48.3	0.1	1.3	2.5
Nicaragua	1.3			0.4		3.4	1.4	0.3		0.1	3.1	3.3
Panama	1.5			0.4		2.6	1.5	0.5		0.2	2.2	2.4
Paraguay	0.9			0.2		3.3	0.7	0.3		0.1	21.4	4.4
Peru	1.1			0.2		2.1	0.8	0.5		0.2	1.6	1.6
Portugal	1.4			0.2		2.0	1.2	0.7		0.4	1.5	1.4
SaudiArabia	1.2			0.1		1.6	0.8	0.8		6.6	1.0	0.8
Singapore	2.0			0.3		1.9	1.6	1.3		3.8	1.3	1.3
SouthAfrica	2.6			0.9		2.4	2.5	0.6		0.2	1.7	1.9
Spain	2.3			0.5		2.2	2.1	1.2		0.5	1.5	1.6
SyrianArabRepublic	1.6					1.9	1.3	1.0		0.8	1.3	1.3
Tunisia	1.5			0.2		2.1	1.3	0.8		0.4	1.5	1.5
Turkey	1.5			0.3		2.4	1.4	0.6		0.2	1.9	2.0
UnitedArabEmirates				0.1								
UnitedKingdom	21.4	93.3	25.8	9.2	83.0	2.3	24.5	46.7	51.7	0.7	1.6	1.9
UnitedStates	3.3			1.1		2.4	2.6	0.6		0.2	1.6	1.9
Venezuela	1.7			0.2		1.9	1.4	1.1		0.8	1.3	1.3

Table G.6: Robustness R6

	AUT	BEL	CAN	CHE	DNK	FIN	FRA	ITA	LUX	NLD	NOR	SWE
Algeria	0.6						0.2					
Argentina	0.9					0.1	0.3	0.1				
Australia	5.0	51.6			57.5	2.9	21.4	22.9	42.7	25.5		
Bahrain						0.1	0.2	23.7		55.0		
Bolivia	0.6					0.1	0.3					
Brazil	0.5					0.1	0.3					
Cameroon	0.4						0.3					
Chile	0.6		56.2			43.6	0.7	0.1			17.1	26.6
Colombia	0.7					0.1	0.4					
CostaRica	0.5					0.1	0.3					
CotedIvoire	0.5					0.1	0.4					
Cuba	0.6					0.1	0.4					
Ecuador	0.4						0.2					
Egypt	0.6						0.2					
ElSalvador	0.6					10.2	0.6				4.5	6.7
Honduras	0.4					0.1	0.3					
HongKongSAR	0.9					0.1	0.3	0.2				
IranRep	0.7						0.2					
Israel	1.0					0.1	0.3	0.2				
Japan	52.7	4.2				0.1	37.5	49.8		7.2		
KoreaRep	0.6					0.1	0.2					
Malaysia	0.6					0.1	0.3					
Mexico	0.8					0.1	0.4					
Morocco	0.5						0.2					
NewZealand	9.6	38.4	43.8	85.8	39.5	21.8	20.8	0.1	56.9		45.5	46.1
Nicaragua	0.6					0.1	0.5					
Nigeria	0.4						0.3					
Panama	0.7					0.1	0.4					
Paraguay	0.4					0.1	0.3					
Peru	0.5					0.1	0.3					
Philippines	0.5					0.1	0.3					
SaudiArabia	0.7					0.1	0.4					
Senegal	0.4					0.1	0.3					
Singapore	0.9					0.1	0.3					
SouthAfrica	1.1					0.1	0.5	0.1				
SyrianArabRepublic	0.6						0.2					
Thailand	0.4					0.1	0.3					
Tunisia	0.6					0.1	0.3					
UnitedArabEmirates	10.8	5.8		14.2	3.0	19.0	8.2	2.0	0.5	12.1	33.0	20.6
Venezuela	0.7					0.1	0.3					
Zimbabwe	0.5					0.1	0.4					

Table G.7: Robustness R7

	AUT	BEL	CAN	CHE	DNK	FIN	FRA	ITA	LUX	NLD	NOR	SWE
Cyprus								31.7		7.9		
Greece										21.2		
Iceland	45.0					0.5	19.8	21.2		52.9	58.5	6.8
Ireland						51.5						
Portugal										18.0		
Turkey	35.1			8.3			13.2	4.2				
UnitedKingdom	19.9	64.5		91.7	44.1		67.0	42.9	73.5		38.8	82.8
UnitedStates		35.5	100.0		55.9	48.0			26.5		2.6	10.3

Table G.8: Robustness R8

	AUT	BEL	CAN	CHE	DNK	FIN	FRA	ITA	LUX	NLD	NOR	SWE
Algeria	0.9					0.2	0.7	0.9		0.3		
Argentina	1.6					0.4	1.3	1.6		0.2		
Australia	1.5					3.1	2.6	0.9		0.1		
Bahrain	14.1					0.3	1.5	11.5		19.4		
Brazil	0.8					0.3	0.7	0.4		0.1		
Chile	0.7		54.5			18.1	1.3	0.4		0.1		7.7
Colombia	1.1					0.4	0.9	0.6		0.1		
CostaRica	0.8					0.3	0.7	0.4		0.1		
CotedIvoire	0.9					0.4	0.9	0.4				
Cuba	0.8					0.4	0.8	0.5		0.1		
Cyprus	1.4					0.4	1.1	1.4		0.6		
Ecuador	0.7					0.2	0.5	0.4		0.1		
ElSalvador	0.9					0.5	0.5	0.4				
Greece	1.3					0.6	1.1	0.9		0.3		
HongKongSAR	1.6					0.6	1.3	1.9		34.7		
Iceland	1.2					28.4	1.6	0.5		0.1	43.8	45.5
IranRep	1.0					0.3	0.8	1.1		0.3		
Ireland	1.3					0.9	1.4	0.7		0.1		
Israel	1.6					0.7	1.4	1.4		0.3		
Japan	4.3					0.6	2.3	2.3		0.2		
KoreaRep	0.9					0.3	0.6	0.7		0.2		
Malaysia	0.9					0.3	0.7	0.6		0.1		
Malta	1.1					0.4	0.9	1.1		0.3		
Mexico	1.2					0.4	1.0	0.9		0.2		
NewZealand	19.9	7.6	45.5	85.7	23.3	0.9	23.4		44.4	0.2	22.5	23.7
Panama	1.0					0.4	1.0	0.5		0.1		
Peru	0.7					0.4	0.7	0.4		0.1		
Portugal	1.0					0.4	0.8	0.6		0.1		
SaudiArabia	0.9					0.4	0.6	0.6		0.2		
Singapore	1.1					0.5	0.9	1.0		0.3		
SouthAfrica	1.4					0.7	1.4	0.9		0.2		
Spain	1.6					0.7	1.4	1.2		0.3		
Tunisia	0.9					0.3	0.7	0.6		0.1		
Turkey	1.0					0.4	0.9	0.6		0.1		
UnitedArabEmirates	12.5	0.1		13.3	3.4		9.3	4.9	0.1	16.6	13.3	
UnitedKingdom	13.2	50.2				1.1	19.0	55.9		23.6		
UnitedStates	1.7	42.0		1.0	73.3	35.0	12.7	0.1	55.5		20.4	22.9
Venezuela	1.0					0.3	0.7	0.8		0.2		

Table G.9: Robustness R9

	AUT	BEL	CAN	CHE	DNK	FIN	FRA	ITA	LUX	NLD	NOR	SWE
Algeria						0.2	0.1	0.1				
Argentina						0.4	0.2	0.1				
Australia							0.3	0.2				
Bahrain	38.2					0.4		49.1		35.2		
Botswana						0.3	0.1	0.1				
Brazil						0.3	0.1	0.1				
Chile			27.6			18.6	0.1					7.7
Colombia						0.3	0.1	0.1				
CostaRica						0.3	0.1	0.1				
CotedIvoire						0.3	0.1	0.1				
Cuba						0.4	0.1	0.1				
Cyprus						0.4	0.2	0.2				
Ecuador						0.2	0.1	0.1				
ElSalvador						0.4						
Greece						0.6	0.2	0.2				
Guatemala						0.2	0.1	0.1				
HongKongSAR						0.5	0.2	0.2				
Iceland	25.3	0.3			9.0	28.1	17.1	3.9	1.2	35.6	78.3	45.6
IranRep						0.2	0.1	0.1				
Ireland						0.9	0.2	0.2				
Israel						0.6	0.2	0.1				
Japan	3.0					0.6	22.5	0.6				
KoreaRep						0.3	0.1	0.1				
Malaysia						0.3	0.1	0.1				
Malta						0.4	0.2	0.1				
Mexico						0.3	0.2	0.1				
NewZealand		7.5	32.3	85.7	19.4	0.6	1.1	0.4	43.1		8.6	23.7
Oman						0.2	0.1	0.1				
Panama						0.3	0.1	0.1				
Peru						0.4	0.1	0.1				
Portugal						0.4	0.2	0.1				
SaudiArabia						0.4	0.2	1.5		15.7		
Singapore						0.5	0.2	0.2				
SouthAfrica						0.5	0.1	0.1				
Spain						0.7	0.3	0.2				
Tunisia						0.3	0.1	0.1				
Turkey						0.3	0.1	0.1				
UnitedArabEmirates				13.4								
UnitedKingdom	33.5	50.2				0.9	54.2	8.5				
UnitedStates		42.0	40.1	0.9	71.6	38.3	0.1	32.4	55.7	13.4	13.1	23.0
Uruguay						0.3	0.1	0.1				
Venezuela						0.3	0.1	0.1				

Table G.10: Robustness R10

	AUT	BEL	CAN	CHE	DNK	FIN	FRA	ITA	LUX	NLD	NOR	SWE
Algeria						0.1	0.1	0.1				
Argentina						0.2	0.2	0.1				
Australia		29.0	37.0		35.0	58.9	0.4	1.2	21.4			14.7
Bahrain	37.9					0.1	0.4	23.3		23.6		
Brazil						0.1	0.1	0.1				
Chile			29.0			1.8						
Colombia						0.1	0.2	0.1				
CostaRica						0.1	0.1	0.1				
CotedIvoire						0.2	0.1					
Cuba						0.1	0.1	0.1				
Cyprus						0.1	0.2	0.2				
Ecuador							0.1	0.1				
ElSalvador		9.3	34.0	58.6	9.8	8.8	0.3		24.4			9.1
Greece						0.2	0.2	0.3				
Honduras						0.1	0.1					
HongKongSAR						0.2	0.1	0.1				
Iceland	25.4					25.8	17.0	4.4		37.0	70.1	23.6
IranRep						0.1	0.1	0.1				
Ireland						0.3	0.2	0.2				
Israel						0.2	0.1	0.1				
Japan	3.5					0.3	22.6	0.4				
KoreaRep						0.1	0.1	0.1				
Malaysia						0.1	0.1	0.1				
Malta						0.1	0.1	0.1				
Mexico						0.1	0.1	0.1				
Morocco							0.1	0.1				
Nicaragua						0.3	0.2					
Panama						0.1	0.1	0.1				
Paraguay						0.1	0.1					
Peru						0.1	0.1	0.1				
Philippines						0.1	0.1					
Portugal						0.1	0.1	0.1				
SaudiArabia						0.1	0.1	7.7		18.4		
Singapore						0.1	0.1	0.2				
SouthAfrica						0.3	0.1	0.1				
Spain						0.2	0.2	0.2				
SyrianArabRepublic						0.1	0.1	0.1				
Tunisia						0.1	0.1	0.1				
Turkey						0.1	0.1	0.1				
UnitedArabEmirates				6.6								
UnitedKingdom	33.2	61.7		34.8	55.2		55.2	59.7	54.3	21.0	29.9	52.6
Venezuela						0.1	0.1	0.1				

G.2 Weights used for constructing synthetic controls. NO_x. 1987

Table G.11: Robustness R1

	AUT	BEL	CAN	CHE	DNK	ESP	FIN	FRA	GBR	GRC	IRL	ITA	LUX	NLD	NOR	SWE	USA
Argentina						0.2											
Australia		17.0	27.3			8.2			10.1		8.0	3.6	56.1				24.5
Cameroon								8.7									
Cuba										3.0							
HongKongSAR	19.9			16.0	31.1		1.8		17.2		14.0	45.0		21.3	19.8	16.4	
Japan	73.3	73.3		72.9	22.0	33.5	0.1	77.4	47.0	32.3	28.8	51.3		63.6	47.8	58.6	
Malta										48.8	8.2						
NewZealand		8.6					39.6	6.6		6.0	24.4						
Oman										9.9							
Portugal						47.5	28.9		7.8		16.7						
UnitedArabEmirates	6.8	1.1	26.7	11.1	46.9		29.6	7.4	17.9				20.8	15.1	32.4	25.0	38.5
Uruguay			46.0			10.6							23.2				37.1

Table G.12: Robustness R2

	AUT	BEL	CAN	CHE	DNK	ESP	FIN	FRA	GBR	GRC	IRL	ITA	LUX	NLD	NOR	SWE	USA
Algeria		0.1				1.1	0.5	0.1	0.4	0.7	0.3						
Argentina		0.4				0.7	0.3	0.2	0.3	0.5	0.2						
Australia		0.5	24.8			1.6	0.6	0.1	0.5	2.2	10.9	3.6	50.3				21.2
Bahrain		0.1				3.2	0.9	0.1	0.7	3.8	1.4						
Brazil		0.2				0.9	0.4	0.1	0.5	0.6	0.2						
Cameroon		0.7				0.5	0.2	0.2	0.2	0.4	0.1						
Chile		0.2				0.9	0.3	0.1	0.4	0.6	0.2						
Colombia		0.1				0.9	0.3	0.1	0.4	0.6	0.2						
CostaRica		0.2				0.8	0.3	0.1	0.4	0.5	0.2						
Cuba		0.1				1.0	0.4	0.1	0.5	0.6	0.3						
Cyprus		0.1				2.6	2.3	0.1	0.9	2.0	0.9						
Ecuador		0.1				1.0	0.5	0.1	0.4	0.7	0.3						
Egypt		0.1				1.0	0.5	0.1	0.4	0.7	0.2						
ElSalvador		0.2				0.8	0.3	0.1	0.4	0.5	0.2						
Honduras		0.1				0.8	0.3	0.1	0.4	0.5	0.2						
HongKongSAR	20.0	0.1		15.6	31.0	9.0	24.4	0.1	11.5	52.5	37.3	45.0		21.3	19.5	16.5	
IranRep		0.1				0.9	0.4	0.1	0.4	0.6	0.2						
Israel		0.1				2.1	1.0	0.1	0.8	1.4	0.7						
Japan	73.2	62.1		72.9	21.8	43.4	19.8	78.9	44.7	12.4	36.6	51.3		63.6	47.7	58.4	
Jordan						1.1	0.6	0.1	0.4	0.7	0.3						
Malaysia		0.1				0.9	0.4	0.1	0.4	0.5	0.2						
Malta		0.1				1.8	1.3	0.1	0.4	1.1	0.6						
Mexico		0.1				1.3	0.6	0.1	0.6	0.8	0.3						
Morocco		0.1				1.0	0.4	0.1	0.4	0.6	0.2						
NewZealand		0.3				3.1	1.3	0.2	1.2	3.4	2.6						
Nicaragua		0.1				0.8	0.3	0.1	0.4	0.5	0.2						
Oman		0.1				1.6	0.8	0.1	0.7	1.0	0.5						
Panama		0.3				0.7	0.3	0.1	0.3	0.5	0.2						
Peru		0.2				0.8	0.3	0.1	0.4	0.5	0.2						
Philippines		0.6				0.5	0.2	0.2	0.1	0.4	0.1						
Portugal		0.1				2.1	1.7	0.1	0.7	1.4	0.7						
SaudiArabia		0.1				2.1	0.7	0.1	0.6	1.0	0.7						
Senegal						0.7				0.4	0.2						
Singapore		0.1				2.3	1.1	0.1	0.8	1.6	0.8						
SouthAfrica		0.7				0.6	0.3	0.2	0.3	0.4	0.2						
SyrianArabRepublic		0.1				0.9	0.4	0.1	0.4	0.6	0.2						
Thailand		0.1				0.9	0.4	0.1	0.4	0.6	0.2						
Tunisia		0.1				1.0	0.4	0.1	0.4	0.6	0.2						
Turkey		0.1				1.1	0.4	0.1	0.5	0.7	0.3						
UnitedArabEmirates	6.8	18.7	28.7	11.5	47.1		33.8	9.1	25.6				25.5	15.1	32.8	25.2	41.1
Uruguay		12.1	46.5			0.4	0.3	7.8		0.4	0.2		24.2				37.7
Venezuela		0.2				0.9	0.4	0.1	0.5	0.6	0.2						

Table G.13: Robustness R3

	AUT	BEL	CAN	CHE	DNK	ESP	FIN	FRA	GBR	GRC	IRL	ITA	LUX	NLD	NOR	SWE	USA
Algeria	1.3	0.6		1.6	1.1	1.9	1.5	0.7	1.1	2.4	2.0	2.3		1.3	1.3	1.1	
Argentina	3.3	2.7		3.3	1.0	3.0	1.3	2.6	2.5	1.2	2.0	2.2		3.0	2.3	2.6	
Australia	7.6	24.2	58.4	5.2	3.5	6.5	4.0	12.5	14.6	1.1	12.3	5.8	79.1	17.0	7.7	19.4	69.9
Bahrain	1.6	0.9		1.8	4.5	2.2	4.3	0.9	2.2	5.5	4.0	3.3		1.8	2.8	2.0	
Brazil	2.0	1.3		2.2	1.3	2.4	1.6	1.3	1.9	1.6	2.2	2.3		2.2	2.0	1.9	
Cameroon	6.6	5.1		5.5	0.8	3.4	1.1	8.5	3.4	1.0	1.8	2.1		4.0	2.5	3.7	
Chile	2.0	1.3		2.2	0.9	2.4	1.2	1.3	1.7	1.4	1.9	2.2		2.1	1.7	1.7	
Colombia	1.8	1.1		2.0	1.0	2.2	1.3	1.1	1.5	1.6	1.9	2.2		1.8	1.6	1.5	
CostaRica	2.1	1.4		2.3	0.8	2.4	1.1	1.3	1.7	1.4	1.8	2.1		2.1	1.6	1.7	
Cuba	1.7	0.9		1.9	1.7	2.2	2.0	1.0	1.7	2.0	2.4	2.5		1.8	1.9	1.6	
Cyprus	1.5	0.7		1.7	1.6	2.0	1.9	0.8	1.4	2.3	2.3	2.4		1.5	1.7	1.3	
Ecuador	1.4	0.7		1.7	1.3	1.9	1.6	0.7	1.3	2.3	2.1	2.3		1.4	1.5	1.2	
Egypt	1.2	0.5		1.5	1.2	1.8	1.5	0.6	1.0	2.9	2.0	2.3		1.2	1.2	0.9	
ElSalvador	2.0	1.3		2.2	0.7	2.4	1.0	1.2	1.5	1.4	1.7	2.1		2.0	1.5	1.6	
Honduras	1.8	1.1		2.0	0.9	2.2	1.2	1.1	1.5	1.6	1.8	2.2		1.8	1.5	1.5	
HongKongSAR	1.3	0.6		1.6	1.9	1.9	2.2	0.6	1.3	3.2	2.5	2.5		1.4	1.6	1.2	
IranRep	1.6	0.8		1.8	1.1	2.1	1.4	0.9	1.4	1.9	2.0	2.3		1.6	1.5	1.3	
Israel	1.6	0.8		1.8	2.0	2.1	2.3	0.9	1.7	2.3	2.6	2.6		1.7	2.0	1.6	
Japan	2.7	2.1		2.9	1.1	2.8	1.4	2.0	2.3	1.3	2.0	2.2		2.7	2.2	2.4	
Jordan	1.1	0.5		1.4	2.2	1.8	2.4	0.5	1.2	5.4	2.7	2.6		1.2	1.5	1.0	
Malaysia	1.9	1.1		2.1	1.3	2.3	1.6	1.2	1.8	1.7	2.2	2.4		2.0	1.9	1.8	
Malta	1.1	0.5		1.4	4.3	1.8	3.2	0.5	1.2	16.4	3.4	2.8		1.2	1.6	1.1	
Mexico	1.5	0.8		1.8	1.4	2.1	1.7	0.9	1.4	2.1	2.2	2.4		1.6	1.6	1.4	
Morocco	1.5	0.7		1.7	0.9	2.0	1.2	0.8	1.2	1.9	1.8	2.2		1.5	1.3	1.2	
NewZealand	2.1	1.4		2.3	1.8	2.5	2.1	1.4	2.3	1.7	2.6	2.5		2.4	2.5	2.3	
Nicaragua	1.9	1.2		2.1	0.9	2.3	1.2	1.2	1.6	1.5	1.8	2.2		1.9	1.6	1.6	
Oman	1.5	0.8		1.8	1.5	2.0	1.8	0.8	1.5	2.2	2.3	2.4		1.6	1.7	1.4	
Panama	2.5	1.8		2.7	0.8	2.7	1.1	1.7	1.9	1.2	1.8	2.1		2.5	1.8	2.0	
Peru	2.3	1.6		2.5	0.8	2.6	1.0	1.5	1.8	1.2	1.7	2.1		2.3	1.7	1.9	
Philippines	5.1	3.1		4.8	0.6	3.2	0.9	5.5	2.5	0.9	1.6	2.0		3.4	2.1	2.8	
Portugal	1.4	0.7		1.7	1.4	1.9	1.7	0.7	1.3	2.4	2.2	2.4		1.4	1.5	1.2	
SaudiArabia	1.6	0.8		1.8	4.8	2.2	4.5	0.9	2.2	7.7	4.1	3.3		1.8	2.8	2.0	
Senegal						2.7				1.1	1.7	2.1					
Singapore	1.5	0.8		1.8	2.2	2.1	2.5	0.9	1.7	2.5	2.7	2.6		1.7	2.0	1.6	
SouthAfrica	6.3	6.2		5.0	1.3	3.6	1.6	7.1	4.0	1.1	2.3	2.3		4.3	3.1	4.3	
SyrianArabRepublic	1.4	0.7		1.7	1.3	1.9	1.6	0.7	1.3	2.3	2.1	2.3		1.4	1.5	1.2	
Thailand	1.5	0.8		1.8	1.1	2.0	1.4	0.8	1.3	2.0	1.9	2.3		1.5	1.4	1.2	
Tunisia	1.5	0.7		1.7	1.0	2.0	1.3	0.8	1.2	2.0	1.9	2.2		1.5	1.4	1.2	
Turkey	1.7	0.9		1.9	1.0	2.1	1.4	1.0	1.4	1.8	1.9	2.2		1.7	1.5	1.4	
UnitedArabEmirates	3.1	0.5		2.9	39.3		29.2	2.0	14.5					2.3	20.6	8.9	
Uruguay	9.9	24.9	41.6	7.3	0.7	3.8	0.9	27.5	4.8	0.9	1.7	2.1	20.9	6.1	2.5	6.1	30.1
Venezuela	2.4	1.7		2.5	1.3	2.7	1.6	1.6	2.2	1.4	2.2	2.3		2.5	2.2	2.3	

Table G.14: Robustness R4

	AUT	BEL	CAN	CHE	DNK	ESP	FIN	FRA	GBR	GRC	IRL	ITA	LUX	NLD	NOR	SWE	USA
Algeria	0.4	0.2			0.5	1.0	1.3	0.1	0.5	0.7	0.3			0.3		0.2	
Argentina	0.5	0.3			0.6	0.7	1.1	0.2	1.4	0.5	0.2			1.0		0.6	
Australia	1.4	0.8			1.0	4.8	1.7	0.5	2.1	28.5	37.6	86.6		1.1		1.1	
Bahrain	0.6	0.3			0.5	3.5	1.5	0.2	0.6	7.3				0.5		0.2	
Brazil	0.6	0.3			0.6	1.0	1.2	0.2	0.9	0.6	0.3			0.6		0.4	
Cameroon	0.3	0.1			0.5	0.5	0.9	0.2	1.2	0.3	0.2			1.1		0.4	
Chile	0.6	0.3			0.6	0.9	1.1	0.2	1.0	0.5	0.3			0.6		0.5	
Colombia	0.5	0.3			0.5	0.9	1.1	0.2	0.8	0.6	0.3			0.5		0.4	
CostaRica	0.5	0.3			0.6	0.9	1.1	0.2	1.0	0.5	0.2			0.6		0.5	
Cuba	0.5	0.2			0.6	1.1	1.3	0.1	0.6	0.7	0.3			0.4		0.3	
Cyprus	0.7	0.3			0.7	2.5	1.6	0.2	0.8	1.8	1.1			0.5		0.4	
Ecuador	0.4	0.2			0.5	1.0	1.2	0.1	0.5	0.7	0.3			0.4		0.2	
Egypt	0.3	0.1			0.5	0.9	1.2	0.1	0.4	0.6	0.3			0.3		0.2	
ElSalvador	0.5	0.3			0.5	0.8	1.0	0.2	0.9	0.5	0.2			0.6		0.4	
Honduras	0.5	0.2			0.5	0.9	1.0	0.2	0.8	0.5	0.2			0.5		0.4	
HongKongSAR	0.7	0.3			0.5	5.8	1.5	0.2	0.6	29.8	27.3	13.4		0.5		0.3	
IranRep	0.4	0.2			0.5	0.9	1.2	0.1	0.6	0.6	0.3			0.4		0.3	
Israel	0.7	0.3			0.7	2.2	1.6	0.2	0.8	1.5	0.8			0.5		0.3	
Japan	59.6	78.3	89.4	53.1	1.4	41.4	1.8	87.0	21.9		18.4		99.1	42.3	31.6	43.9	89.6
Jordan	0.3	0.1			0.4	1.0	1.4	0.1	0.4	0.6	0.3			0.3		0.1	
Malaysia	0.5	0.2			0.5	1.0	1.1	0.2	0.8	0.6	0.3			0.5		0.4	
Malta	0.4	0.2			0.4	2.7	1.5	0.1	0.4	7.4	0.2			0.3		0.2	
Mexico	0.5	0.2			0.6	1.3	1.4	0.2	0.7	0.8	0.4			0.5		0.3	
Morocco	0.4	0.2			0.5	0.9	1.1	0.1	0.6	0.6	0.3			0.4		0.3	
NewZealand	1.2	0.7			1.0	3.7	1.7	0.4	1.9	1.3	4.2			1.0		0.9	
Nicaragua	0.5	0.3			0.5	0.8	1.0	0.2	0.8	0.5	0.2			0.5		0.4	
Oman	0.6	0.3			0.7	1.7	1.5	0.2	0.8	1.0	0.5			0.5		0.3	
Panama	0.5	0.3			0.6	0.7	1.0	0.2	1.2	0.5	0.2			0.8		0.6	
Peru	0.5	0.3			0.5	0.8	1.0	0.2	1.1	0.5	0.2			0.7		0.5	
Philippines	0.2				0.5	0.4	0.8	0.2	1.0	0.3	0.2			1.2		0.3	
Portugal	0.6	0.3			0.7	2.1	1.5	0.2	0.7	1.4	0.7			0.5		0.3	
SaudiArabia	0.5	0.2			0.5	2.5	1.5	0.1	0.5	2.6	0.6			0.4		0.2	
Senegal						0.6				0.4	0.2						
Singapore	0.7	0.3			0.6	2.4	1.6	0.2	0.7	1.8	0.9			0.5		0.3	
SouthAfrica	0.5	0.3			0.6	0.6	1.1	0.2	1.5	0.5	0.2			1.1		0.7	
SyrianArabRepublic	0.4	0.2			0.5	0.9	1.2	0.1	0.5	0.6	0.3			0.3		0.2	
Thailand	0.4	0.2			0.5	0.9	1.1	0.1	0.6	0.6	0.3			0.4		0.3	
Tunisia	0.4	0.2			0.5	1.0	1.2	0.1	0.6	0.6	0.3			0.4		0.3	
Turkey	0.5	0.3			0.6	1.1	1.2	0.2	0.8	0.7	0.3			0.5		0.4	
UnitedArabEmirates	19.7	9.1		46.9	76.5		49.5	4.2	40.6					34.1	68.4	40.7	10.4
Uruguay		2.4	10.6		0.6	0.3	0.9	1.9	5.3	0.4	0.2		0.9	1.7		0.8	
Venezuela	0.6	0.3			0.6	1.0	1.2	0.2	1.1	0.6	0.3			0.7		0.5	

Table G.15: Robustness R5

	AUT	BEL	CAN	CHE	DNK	ESP	FIN	FRA	GBR	GRC	IRL	ITA	LUX	NLD	NOR	SWE	USA
Algeria	1.8	1.3	0.6	1.4	3.8	2.4	0.6	1.0	1.7	2.8	3.6	2.4			0.9	0.5	1.0
Argentina	4.1	4.7		1.6	1.7	4.5	0.3	2.2	2.9	0.8	3.4	1.5			0.4	0.3	4.3
Australia	2.4	2.1	1.0	1.6	2.4	2.8	0.4	1.5	1.8	1.7	3.0	1.6			0.6	0.4	1.5
Bahrain	1.5	1.0	0.4	1.4	3.1	2.1	0.9	0.8	1.3	2.5	2.8	1.1			1.3	0.6	0.5
Brazil	1.5	1.1	0.7	2.3	0.6	1.6	0.6	1.8	0.4	0.3	1.4	0.3			1.4	0.9	0.3
Cameroon	1.2	0.8		13.3	0.3	1.1	0.4	9.9	0.2	0.2	1.2	0.3			0.9	4.4	
Chile	1.8	1.5	1.0	2.0	0.7	1.9	0.5	1.9	0.5	0.4	1.7	0.5			0.8	0.6	0.4
Colombia	1.7	1.3	0.8	2.0	0.7	1.9	0.6	1.7	0.5	0.4	1.7	0.5			0.9	0.7	0.4
CostaRica	1.7	1.4	0.9	2.3	0.6	1.7	0.6	2.1	0.4	0.3	1.5	0.4			0.9	0.8	0.3
Cuba	1.4	1.0	0.6	1.8	0.8	1.7	0.9	1.3	0.5	0.5	1.6	0.4			1.7	0.8	0.3
Cyprus	1.9	1.4	0.6	1.4	3.8	2.5	0.6	1.0	1.8	2.8	3.6	2.6			0.9	0.5	1.1
Ecuador	1.3	0.9	0.5	1.7	0.9	1.7	1.2	1.1	0.6	0.6	1.7	0.4			1.9	0.8	0.3
Egypt	1.4	1.0	0.5	1.4	1.8	2.0	1.0	0.9	1.0	1.3	2.2	0.7			1.4	0.7	0.4
ElSalvador	1.3	0.9	0.7	3.1	0.4	1.3	0.1	2.3	0.3	0.2	1.3	0.3			2.3		0.2
Honduras	1.2	0.8	0.6	17.4	0.4	1.3	50.0	1.9	0.3	0.2	1.2	0.3			50.8	73.7	0.2
HongKongSAR	1.5	1.0	0.5	1.4	2.6	2.1	0.9	0.8	1.2	2.0	2.6	1.0			1.3	0.6	0.5
IranRep	2.2	1.8	0.8	1.5	17.4	2.9	0.5	1.3	32.1	7.8	7.0	38.4		67.1	0.7	0.5	9.5
Israel	1.8	1.3	0.6	1.5	2.5	2.3	0.6	1.1	1.4	1.9	2.7	1.2			1.0	0.6	0.8
Japan	2.8	3.0	1.3	1.7	1.2	3.1	0.4	2.1	1.2	0.6	2.1	0.8			0.5	0.3	1.4
Jordan	1.3	0.8	0.3	1.3	3.9	1.9	2.3	0.7	1.1	3.2	2.6	0.7			2.0	0.7	0.4
Malaysia	2.0	1.7	0.9	1.7	1.2	2.3	0.5	1.5	0.9	0.8	2.0	0.7			0.7	0.5	0.7
Malta	1.2	0.7	0.3	1.3	10.4	1.9	25.3	0.6	1.0	12.7	2.6	0.6			7.1	0.7	0.3
Mexico	1.7	1.3	0.7	1.6	1.4	2.1	0.7	1.2	0.9	0.9	2.1	0.7			1.0	0.6	0.5
Morocco	2.0	1.6	0.8	1.6	1.9	2.4	0.5	1.3	1.3	1.4	2.5	1.1			0.8	0.5	0.9
NewZealand	1.7	1.3	0.8	1.9	0.8	1.9	0.6	1.6	0.5	0.5	1.7	0.5			1.0	0.7	0.4
Nicaragua	1.4	1.0	0.7	2.6	0.5	1.4	0.3	2.0	0.3	0.3	1.3	0.3			2.1	0.7	0.2
Oman	2.0	1.5	0.7	1.5	8.2	2.7	0.5	1.1	7.5	36.5	6.4	15.5			0.8	0.5	1.4
Panama	2.2	2.0	1.2	2.1	0.6	2.1	0.4	2.3	0.5	0.4	1.7	0.5			0.6	0.5	0.5
Peru	2.1	1.8	1.2	2.1	0.6	2.0	0.5	2.2	0.5	0.4	1.6	0.5			0.6	0.5	0.5
Philippines	2.7	3.0	2.2	3.8	0.4	1.4	0.4	15.1	0.2	0.2	1.4	0.4			0.5	0.2	0.2
Portugal	1.4	1.0	0.5	1.6	1.1	1.9	1.0	1.1	0.7	0.7	1.9	0.5			1.5	0.7	0.4
SaudiArabia	1.5	0.9	0.4	1.4	3.2	2.1	1.0	0.8	1.2	2.5	2.8	1.1			1.4	0.6	0.5
Senegal						1.6				0.3	1.4	0.4					
Singapore	1.8	1.2	0.6	1.4	3.2	2.3	0.6	1.0	1.5	2.4	3.1	1.7			1.0	0.6	0.8
SouthAfrica	21.4	22.4	38.9	1.6	3.3	15.8	0.3	2.3	23.1	4.0	5.1	13.8	95.5	32.9	0.4	0.2	57.1
SyrianArabRepublic	1.8	1.3	0.6	1.5	2.5	2.3	0.6	1.1	1.4	1.9	2.8	1.3			0.9	0.6	0.8
Thailand	1.3	0.9	0.5	2.0	0.6	1.5	0.7	1.4	0.4	0.4	1.4	0.3			3.2	0.8	0.2
Tunisia	1.7	1.3	0.7	1.6	1.2	2.1	0.6	1.3	0.9	0.8	2.0	0.6			1.0	0.6	0.5
Turkey	1.7	1.4	0.8	1.8	1.0	2.0	0.6	1.4	0.7	0.6	1.8	0.6			0.9	0.6	0.5
UnitedArabEmirates	1.9	1.4	0.6	1.4	5.1		0.6	1.1	2.4						0.9	0.5	1.3
Uruguay	7.9	20.2	33.5	1.9	0.5	1.9	0.4	20.8	0.3	0.3	1.6	0.5	4.5		0.4	0.1	5.4
Venezuela	3.0	3.0	1.2	1.6	2.7	3.4	0.4	1.8	2.7	1.8	3.8	3.0			0.5	0.4	2.9

Table G.16: Robustness R6

	AUT	BEL	CAN	CHE	DNK	ESP	FIN	FRA	GBR	GRC	IRL	ITA	LUX	NLD	NOR	SWE	USA
Algeria						1.3			0.2	1.7	0.7						
Argentina						0.8			0.1	0.4	0.4						
Australia			28.0			0.5			0.1	0.2	0.4		58.9				24.1
Bahrain						1.0			0.1	0.3	0.7						
Brazil		0.1				1.0				0.2	0.5						
Cameroon		13.3				0.8		8.1		0.1	0.3						
Chile						1.0				0.2	0.5						
Colombia						1.1				0.2	0.5						
CostaRica		0.1				1.1				0.2	0.5						
Cuba						1.0				0.2	0.6						
Ecuador						1.2				0.3	0.7						
Egypt						1.3				0.4	0.7						
ElSalvador						1.2				0.2	0.5						
Honduras						1.3				0.2	0.7						
HongKongSAR	20.0			16.6	31.3	16.7	23.8		1.9	40.0	29.3	44.8		21.4	20.2	16.4	
IranRep						1.0			0.3	1.1	0.5						
Israel						1.5			0.1	0.8	1.0						
Japan	73.2	64.8		72.2	21.7	39.5		78.9	44.6	10.8	29.9	51.2		63.5	47.3	58.5	
Jordan						1.1					0.7						
Malaysia						0.9				0.3	0.5						
Mexico						1.3			0.1	0.4	0.7						
Morocco						1.2			0.1	0.5	0.6						
NewZealand						2.8	43.4		0.1	0.4	5.8						
Nicaragua						1.1				0.2	0.5						
Nigeria		2.1				2.0	4.9	3.7		0.2	5.2						
Oman						1.5			25.0	32.7	0.9						
Panama		0.1				0.9				0.2	0.4						
Peru		0.1				1.0				0.2	0.4						
Philippines		0.1				0.8				0.1	0.3						
SaudiArabia						0.9			0.1	0.2	0.6						
Senegal		0.1				0.9				0.2	0.4						
Singapore						1.5			0.2	1.1	1.0						
SouthAfrica						0.6			0.1	0.3	0.3						
SyrianArabRepublic						1.1			0.1	0.5	0.6						
Thailand						1.2				0.2	0.6						
Tunisia						1.2				0.3	0.6						
UnitedArabEmirates	6.8	18.6	25.7	11.2	47.1	3.5	27.8	9.3	26.4	4.1	10.7	4.0	17.4	15.1	32.5	25.1	38.4
Uruguay			46.3			0.8				0.2	0.3		23.7				37.5
Venezuela						0.8			0.1	0.5	0.4						

Table G.17: Robustness R7

	AUT	BEL	CAN	CHE	DNK	ESP	FIN	FRA	GBR	GRC	IRL	ITA	LUX	NLD	NOR	SWE	USA
Cyprus	100.0	100.0	92.3	98.9	100.0	93.1	100.0	100.0	100.0	94.0	100.0	100.0	35.2	100.0	100.0	100.0	73.2
Malta										6.0							
Turkey			7.7	1.1		6.9							64.8				26.8

Table G.18: Robustness R8

	AUT	BEL	CAN	CHE	DNK	ESP	FIN	FRA	GBR	GRC	IRL	ITA	LUX	NLD	NOR	SWE	USA
Algeria						1.3			0.2	1.0	0.8						
Argentina						0.8			0.9	0.6	0.4						
Australia			28.2			0.8			0.2	0.6	0.3		60.1				24.6
Bahrain						1.1			0.2	0.7	0.7						
Brazil						1.3				0.5	0.6						
Chile						1.3			0.1	0.6	0.6						
Colombia						1.3			0.1	0.6	0.7						
CostaRica						1.4				0.5	0.7						
Cuba						1.2			0.1	0.5	0.7						
Cyprus						2.3			14.9	57.9	1.5						
Ecuador						1.5				0.6	0.9						
ElSalvador		6.4				5.9	1.4	7.0		0.4	0.2						
HongKongSAR	20.0			16.8	31.4	9.4				11.9	22.7	44.9		21.5	20.4	12.1	
IranRep						1.1			0.4	0.8	0.6						
Israel						1.6			0.2	1.2	1.0						
Japan	73.2	60.3		71.9	21.6	38.7		75.6	42.2	4.5	23.3	51.1		63.4	47.1	50.7	
Jordan						1.1			0.1	0.9	0.7						
Malaysia						1.0			0.1	0.6	0.5						
Malta						1.5			0.1	1.9	1.0						
Mexico						1.4			0.1	0.9	0.8						
NewZealand						4.2	37.2		0.1	0.7	26.6					14.4	
Oman						1.5			11.3	1.2	0.9						
Panama						1.1			0.1	0.5	0.5						
Peru						1.2			0.1	0.5	0.6						
Portugal						5.8	30.8		0.1	1.0	2.4						
SaudiArabia						1.0			0.1	0.6	0.6						
Singapore						1.6			0.3	1.3	1.0						
SouthAfrica						0.6			1.3	0.5	0.4						
Tunisia						1.3			0.1	0.8	0.7						
Turkey						1.5			0.1	0.7	0.8						
UnitedArabEmirates	6.8	20.2	25.3	11.2	47.1	2.7	30.6	10.2	26.0	3.9	5.6	4.0	16.0	15.1	32.6	22.8	37.8
Uruguay		13.1	46.5			0.4		7.2		0.4	0.4		24.0				37.6
Venezuela						1.0			0.5	0.7	0.5						

Table G.19: Robustness R9

	AUT	BEL	CAN	CHE	DNK	ESP	FIN	FRA	GBR	GRC	IRL	ITA	LUX	NLD	NOR	SWE	USA
Algeria						1.0			0.1	0.3	0.2						
Argentina						0.6			0.7	0.3	0.2						
Australia						0.7			0.1	4.2	5.0	3.7					
Bahrain						1.6			0.1	1.6	1.2						
Botswana			16.3			0.1				0.1	0.1		28.1				16.2
Brazil						1.0				0.3	0.2						
Cameroon		11.7				0.7		6.1		0.2	0.1						
Chile						0.8				0.3	0.2						
Colombia						0.9				0.3	0.2						
CostaRica						0.8				0.3	0.2						
CotedIvoire		0.2	10.4			0.4				0.2	0.1						
Cuba						1.0				0.3	0.2						
Cyprus						2.4			8.0	56.0	0.7						
Ecuador						1.2				0.3	0.2						
ElSalvador						1.4				0.2	0.2						
Guatemala		4.1				1.2		6.0		0.2	0.1						
HongKongSAR	19.6			8.7	29.2	12.1	9.0		1.2	20.4	30.5	43.2		20.9	14.1	15.4	
IranRep						0.8			0.3	0.3	0.2						
Israel						1.8			0.1	0.6	0.6						
Japan	73.6	64.4	35.5	79.0	23.5	37.6		78.4	43.2	5.8	23.7	53.1	13.6	64.0	52.6	59.4	35.7
Jordan						1.3			0.1	0.3	0.3						
Malaysia						0.7			0.1	0.3	0.2						
Malta						1.5				1.1	0.5						
Mexico						1.2			0.1	0.4	0.3						
Morocco						0.9			0.1	0.3	0.2						
Namibia						0.3				0.2	0.1						
NewZealand		0.1				11.1	44.3		0.1	0.5	30.7						
Nicaragua		0.1				1.1				0.2	0.2						
Oman						1.4			17.7	0.4	0.4						
Panama						0.7				0.3	0.2						
Paraguay						0.6				0.2	0.1						
Peru						0.7				0.3	0.2						
Portugal						2.9	18.7		0.1	0.6	0.6						
SaudiArabia						1.3			0.1	0.5	0.6						
Singapore						1.9			0.1	0.7	0.7						
SouthAfrica						0.5			0.8	0.2	0.2						1.7
SyrianArabRepublic						0.9				0.3	0.2						
Tunisia						0.9			0.1	0.3	0.2						
Turkey						1.0			0.1	0.3	0.2						
UnitedArabEmirates	6.8	18.7	37.9	12.3	47.3		28.1	9.5	26.2				58.2	15.2	33.3	25.2	46.4
Uruguay						0.6				0.2	0.1						
Venezuela						0.7			0.2	0.3	0.2						

Table G.20: Robustness R10

	AUT	BEL	CAN	CHE	DNK	ESP	FIN	FRA	GBR	GRC	IRL	ITA	LUX	NLD	NOR	SWE	USA
Algeria										0.5							
Argentina										0.3							
Australia			21.4							0.2			44.5				18.5
Bahrain										0.2							
Brazil										0.4							
Cameroon										0.2							
Chile										0.4							
Colombia										0.4							
CostaRica										0.4							
Cuba										0.3							
Cyprus	39.8	16.4		43.6	43.9	57.3		19.5	37.2	74.5	62.0	65.1		39.7	40.3	32.6	
Ecuador										0.4							
Egypt										0.4							
ElSalvador							1.7			0.3							
Honduras										0.4							
HongKongSAR										4.9							
IranRep										0.4							
Israel										0.6							
Jordan										0.1							
Malaysia										0.4							
Mexico										0.5							
Morocco										0.5							
Nicaragua										0.4							
Nigeria										0.4							
Oman										0.7							
Panama										0.3							
Peru										0.3							
Philippines										0.2							
Portugal							46.3			0.5							
SaudiArabia										0.2							
Senegal										0.3							
Singapore										0.6							
SouthAfrica										0.3							
SyrianArabRepublic										0.4							
Thailand										0.4							
Tunisia										0.5							
Turkey										0.5							
UnitedArabEmirates	18.7	29.4	31.6	21.9	50.5	9.5	40.8	22.1	33.5	5.5	16.0	12.3	30.4	25.4	39.7	34.6	43.4
Uruguay	41.5	54.2	47.0	34.5	5.6	33.3	11.2	58.4	29.3	1.5	22.0	22.6	25.1	34.9	20.0	32.8	38.1
Venezuela										0.4							

G.3 Weights used for constructing synthetic controls. VOC. 1990.

Table G.21: Robustness R1

	AUT	BEL	CHE	DNK	ESP	FIN	FRA	GBR	ITA	LUX	NLD	SWE
Australia		3.5	6.4			18.5					3.5	
Bahrain	7.6			3.2	10.5	4.5	5.4	12.1	7.8	26.5		10.3
Cameroon		7.4					6.8					
HongKongSAR					80.1			35.2	31.7	34.8		18.7
Japan	84.2	86.8	78.3	96.8	9.4	69.0	85.8	52.7	60.5	38.6	85.9	70.9
NewZealand	7.0	2.2				8.0	2.0					
Oman	1.2											
SaudiArabia			15.3								10.6	

Table G.22: Robustness R2

	AUT	BEL	CHE	DNK	ESP	FIN	FRA	GBR	ITA	LUX	NLD	SWE
Algeria		0.2					0.2					
Argentina		0.2					0.2					
Australia	0.1	3.3	6.7			16.5	1.2				3.5	
Bahrain	8.7	1.7		3.2	10.5	6.3	6.3	12.1	7.8	26.1		10.3
Brazil		0.2					0.2					
Cameroon		0.2					0.2					
Chile		0.2					0.2					
Colombia		0.2					0.2					
CostaRica		0.2					0.2					
CotedIvoire		0.2					0.1					
Cuba	0.1	0.2					0.3					
Cyprus	0.1	0.2				0.1	0.3					
Ecuador		0.2					0.2					
Egypt		0.2					0.2					
ElSalvador		0.2					0.1					
Honduras		0.2					0.2					
HongKongSAR	2.5	0.1			80.1	0.1	0.2	34.8	31.6	33.0		18.5
IranRep		0.2					0.1					
Ireland						0.2	0.1					
Japan	87.3	86.7	78.3	96.8	9.3	75.7	85.3	53.1	60.6	40.9	86.0	71.2
Jordan		0.2					0.2					
Malaysia		0.2					0.2					
Malta	0.1	0.3					0.3					
Mexico		0.3					0.2					
Morocco		0.2					0.1					
NewZealand	0.1	0.2				0.1	0.3					
Nicaragua		0.2					0.1					
Oman	0.2	0.4					0.5					
Panama		0.2					0.1					
Paraguay		0.2					0.2					
Peru		0.2					0.1					
Philippines		0.2					0.1					
SaudiArabia		0.4	15.0				0.2				10.5	
Singapore		0.3				0.1	0.3					
SouthAfrica		0.2					0.2					
SyrianArabRepublic		0.2					0.2					
Thailand		0.2					0.2					
Tunisia		0.2					0.2					
Turkey		0.2					0.2					
Uruguay		0.2					0.2					
Venezuela		0.3					0.2					

Table G.23: Robustness R3

	AUT	BEL	CHE	DNK	ESP	FIN	FRA	GBR	ITA	LUX	NLD	SWE
Algeria	2.4	2.6	1.8	2.4	0.6	2.1	2.4	1.6	1.9	0.2	2.1	2.0
Argentina	2.0	2.4	2.6	1.9	0.3	2.2	2.2	1.0	1.3	0.1	2.6	1.3
Australia	2.3	2.0	2.1	1.9	0.3	3.3	2.4	1.1	1.4	0.1	2.2	1.5
Bahrain	5.6	1.8	1.3	2.4	0.7	9.9	4.2	2.3	2.3	1.7	1.5	4.9
Brazil	2.1	2.3	2.4	1.9	0.4	2.3	2.2	1.1	1.4	0.1	2.5	1.4
Cameroon	2.4	2.2	1.9	2.1	0.5	2.6	2.4	1.4	1.7	0.2	2.1	1.8
Chile	2.6	2.6	1.5	2.9	1.0	2.0	2.6	2.3	2.7	0.2	2.0	2.6
Colombia	2.4	2.5	1.8	2.4	0.5	2.1	2.4	1.5	1.9	0.2	2.1	1.9
CostaRica	2.2	2.5	2.1	2.2	0.4	2.1	2.3	1.3	1.6	0.1	2.3	1.6
CotedIvoire	2.0	2.3	2.8	1.8	0.3	2.4	2.1	0.9	1.2	0.1	2.7	1.2
Cuba	3.6	2.4	1.2	5.4	39.1	2.2	3.1	18.8	12.8	21.0	1.7	15.7
Cyprus	2.8	2.6	1.4	3.3	2.0	2.1	2.7	3.2	3.6	0.2	1.9	3.2
Ecuador	2.4	2.5	1.8	2.3	0.5	2.1	2.4	1.5	1.9	0.1	2.2	1.9
Egypt	2.4	2.7	1.7	2.6	0.6	2.0	2.5	1.8	2.1	0.2	2.1	2.1
ElSalvador	1.9	2.5	3.3	1.8	0.3	2.1	2.1	0.9	1.2	0.1	3.1	1.1
Honduras	2.3	2.5	1.9	2.3	0.5	2.1	2.4	1.4	1.8	0.1	2.2	1.8
HongKongSAR	2.9	3.0	1.3	4.6	26.3	1.9	2.7	5.0	8.6	0.2	1.9	3.7
IranRep	2.0	2.4	2.7	1.9	0.3	2.2	2.2	1.0	1.3	0.1	2.7	1.3
Ireland	2.3	2.4	1.9	2.2	0.5	2.2	2.4	1.4	1.7	0.1	2.2	1.8
Japan	2.3	2.4	2.0	2.2	0.5	2.2	2.4	1.4	1.7	0.1	2.2	1.7
Jordan	2.3	2.7	2.0	2.3	0.5	2.0	2.3	1.4	1.7	0.1	2.3	1.8
Malaysia	2.1	2.2	2.4	1.9	0.3	2.5	2.2	1.0	1.4	0.1	2.5	1.4
Malta	3.0	2.7	1.3	4.3	10.7	2.0	2.8	5.4	7.2	0.2	1.8	3.9
Mexico	2.5	2.5	1.6	2.5	0.7	2.2	2.5	1.8	2.1	0.2	2.0	2.2
Morocco	2.1	2.7	2.4	2.1	0.4	2.0	2.3	1.2	1.6	0.1	2.6	1.6
NewZealand	2.5	2.4	1.7	2.4	0.6	2.3	2.5	1.6	1.9	0.2	2.1	2.1
Nicaragua	2.0	2.4	2.8	1.9	0.3	2.1	2.1	1.0	1.3	0.1	2.7	1.3
Oman	3.6	2.2	1.2	3.4	2.1	2.7	3.2	16.6	5.1	72.4	1.7	8.7
Panama	1.8	2.4	3.8	1.7	0.2	2.2	2.0	0.7	1.1		3.3	1.0
Paraguay	2.4	2.1	2.0	2.1	0.4	2.8	2.4	1.3	1.6	0.1	2.1	1.7
Peru	1.7	2.4	4.5	1.6	0.2	2.2	2.0	0.7	1.0		4.3	0.9
Philippines	2.0	2.5	2.8	1.9	0.3	2.0	2.1	1.0	1.4	0.1	2.8	1.3
SaudiArabia	1.7	1.9	17.1	1.3	0.1	3.9	1.9	0.5	0.7		8.2	0.7
Singapore	2.5	2.4	1.7	2.4	0.6	2.2	2.5	1.6	1.9	0.2	2.1	2.0
SouthAfrica	2.6	2.5	1.6	2.6	0.8	2.1	2.6	2.0	2.3	0.2	2.0	2.4
SyrianArabRepublic	2.8	2.8	1.3	3.6	3.2	2.0	2.7	3.5	4.3	0.2	1.9	3.2
Thailand	2.6	2.5	1.6	2.6	0.8	2.1	2.5	2.0	2.3	0.2	2.0	2.4
Tunisia	2.2	2.7	2.1	2.2	0.4	2.0	2.3	1.3	1.7	0.1	2.4	1.7
Turkey	2.6	2.6	1.6	2.7	0.8	2.0	2.5	2.1	2.4	0.2	2.0	2.4
Uruguay	2.2	2.6	2.2	2.1	0.4	2.0	2.3	1.3	1.6	0.1	2.4	1.6
Venezuela	2.1	2.2	2.5	1.8	0.3	2.6	2.2	1.0	1.3	0.1	2.5	1.3

Table G.24: Robustness R4

	AUT	BEL	CHE	DNK	ESP	FIN	FRA	GBR	ITA	LUX	NLD	SWE
Algeria	0.2	0.2			0.1	0.1	0.3					
Argentina	0.2	0.3			0.1	0.2	0.3					
Australia		0.9	61.3		0.2	7.2	0.2				36.9	
Bahrain	0.2	0.3			0.2	0.2	0.4					
Brazil	0.2	0.3			0.1	0.2	0.3					
Cameroon	0.1	0.2			0.1	0.1	0.3					
Chile	0.2	0.2			0.1	0.1	0.3					
Colombia	0.2	0.2			0.1	0.1	0.3					
CostaRica	0.2	0.3			0.1	0.1	0.3					
CotedIvoire	0.1	0.3			0.1	0.2	0.3					
Cuba	0.3	0.2				0.1	0.3					
Cyprus	0.4	0.2			0.3	0.1	0.3		0.1			
Ecuador	0.2	0.2			0.1	0.1	0.3					
Egypt	0.2	0.2			0.1	0.1	0.3					
ElSalvador	0.1	0.3			0.1	0.2	0.3					
Honduras	0.2	0.2			0.1	0.1	0.3					
HongKongSAR	2.2	0.2		5.0	82.0	0.1	0.3	41.6	33.9	58.2		25.7
IranRep	0.1	0.3			0.1	0.2	0.3					
Ireland	0.1	0.1			0.3	0.1	0.2		0.1			
Japan	90.5	88.8	38.7	95.0	14.1	84.6	88.3	58.4	65.7	41.8	63.1	74.3
Jordan	0.2	0.2			0.1	0.1	0.3					
Malaysia	0.1	0.3			0.1	0.2	0.3					
Malta	0.4	0.3			0.1	0.1	0.4					
Mexico	0.2	0.3			0.1	0.1	0.4					
Morocco	0.2	0.2			0.1	0.1	0.3					
NewZealand	0.2	0.2			0.2	0.1	0.2					
Nicaragua	0.1	0.2			0.1	0.2	0.3					
Oman	0.4	0.3			0.1	0.1	0.4					
Panama	0.1	0.3			0.1	0.2	0.3					
Paraguay	0.1	0.2			0.1	0.1	0.3					
Peru	0.1	0.3			0.1	0.2	0.3					
Philippines	0.1	0.2			0.1	0.1	0.3					
SaudiArabia	0.1	0.9				2.7	0.4					
Singapore	0.3	0.3			0.2	0.2	0.4					
SouthAfrica	0.2	0.3			0.1	0.1	0.3					
SyrianArabRepublic	0.2	0.2				0.1	0.3					
Thailand	0.2	0.2			0.1	0.1	0.3					
Tunisia	0.2	0.2			0.1	0.1	0.3					
Turkey	0.2	0.2			0.1	0.1	0.3					
Uruguay	0.2	0.3			0.1	0.2	0.3					
Venezuela	0.2	0.3			0.1	0.2	0.3					

Table G.25: Robustness R5

	AUT	BEL	CHE	DNK	ESP	FIN	FRA	GBR	ITA	LUX	NLD	SWE
Algeria	3.0	2.6	1.2	6.3	0.6	1.4	1.4	3.1	4.8	0.6	6.6	0.1
Argentina	2.3	2.4	1.8	1.5	0.4	1.4	1.3	1.1	1.1	0.4	0.9	0.1
Australia	2.4	2.4	1.9	1.8	0.3	1.3	1.2	1.2	1.3	0.3	1.3	0.1
Bahrain	2.9	2.6	1.3	6.4	0.5	1.4	1.3	3.0	4.6	0.5	16.1	0.1
Brazil	1.9	2.3	2.1	0.9	0.4	2.2	1.9	0.7	0.5	0.4	0.4	0.1
Cameroon	1.7	2.2	2.3	0.7	0.5	19.7	17.8	0.6	0.4	0.5	0.2	55.9
Chile	2.3	2.5	1.2	1.2	0.8	1.9	2.3	1.3	0.8	0.9	0.4	0.2
Colombia	2.2	2.4	1.5	1.1	0.6	2.0	2.0	1.0	0.7	0.6	0.4	0.2
CostaRica	2.0	2.3	1.7	1.0	0.5	2.1	2.1	0.9	0.6	0.5	0.4	0.1
CotedIvoire	1.7	2.1	30.0	0.7	0.3	6.1	2.6	0.5	0.3	0.3	0.3	
Cuba	2.2	2.5	1.1	0.8	44.2	2.3	7.3	1.6	0.5	13.4	0.3	39.7
Cyprus	3.6	2.8	0.9	5.6	1.2	1.5	1.6	4.9	6.0	1.6	4.1	0.1
Ecuador	2.3	2.4	1.4	1.3	0.6	1.7	1.7	1.2	0.9	0.6	0.5	0.1
Egypt	2.8	2.6	1.1	2.1	0.7	1.5	1.6	1.9	1.9	0.7	0.9	0.1
ElSalvador	1.8	2.2	2.4	0.8	0.3	3.7	2.3	0.6	0.4	0.3	0.3	0.1
Honduras	1.8	2.2	1.9	0.8	0.6	3.7	6.9	0.7	0.4	0.5	0.2	0.9
HongKongSAR	3.6	2.8	0.9	5.1	26.5	1.5	1.7	20.3	13.0	48.8	1.7	0.1
IranRep	2.6	2.5	1.7	3.7	0.4	1.3	1.2	2.1	2.6	0.4	2.6	0.1
Ireland	2.4	2.5	1.4	1.5	0.5	1.6	1.6	1.3	1.1	0.6	0.6	0.1
Japan	2.4	2.5	1.4	1.5	0.5	1.6	1.5	1.2	1.1	0.5	0.6	0.1
Jordan	2.8	2.6	1.3	3.0	0.5	1.4	1.4	2.2	2.6	0.6	1.8	0.1
Malaysia	2.3	2.4	1.8	1.5	0.4	1.4	1.3	1.1	1.0	0.4	0.9	0.1
Malta	4.3	2.8	0.9	10.3	8.2	1.5	1.7	17.7	19.0	14.4	9.2	0.1
Mexico	2.5	2.5	1.2	1.6	0.6	1.6	1.7	1.5	1.3	0.7	0.6	0.1
Morocco	2.6	2.5	1.4	1.9	0.5	1.4	1.4	1.5	1.6	0.5	1.0	0.1
NewZealand	2.2	2.4	1.5	1.1	0.6	2.0	2.1	1.0	0.7	0.6	0.4	0.2
Nicaragua	1.8	2.2	2.3	0.8	0.4	4.0	2.7	0.6	0.4	0.4	0.3	0.1
Oman	3.7	2.8	0.9	9.3	1.2	1.5	1.6	5.8	9.3	1.7	15.2	0.1
Panama	1.9	2.2	2.5	1.0	0.3	1.7	1.5	0.6	0.5	0.3	0.5	0.1
Paraguay	1.7	2.2	2.3	0.7	0.4	5.3	4.5	0.6	0.4	0.4	0.2	
Peru	1.9	2.2	2.5	1.0	0.2	1.6	1.4	0.6	0.5	0.2	0.5	0.1
Philippines	1.9	2.2	2.1	0.9	0.4	2.6	2.2	0.7	0.5	0.4	0.3	0.1
SaudiArabia	2.1	2.3	9.3	5.6	0.1	0.7	0.9	1.7	4.8	0.1	22.0	
Singapore	3.0	2.6	1.2	4.7	0.6	1.4	1.4	2.8	3.9	0.6	3.5	0.1
SouthAfrica	2.5	2.5	1.2	1.6	0.7	1.6	1.7	1.5	1.2	0.7	0.5	0.1
SyrianArabRepublic	3.3	2.8	0.9	3.6	1.8	1.5	1.7	4.9	4.6	2.6	1.3	0.1
Thailand	2.2	2.4	1.3	1.1	0.7	2.1	2.5	1.1	0.7	0.7	0.3	0.2
Tunisia	2.4	2.5	1.4	1.6	0.5	1.6	1.5	1.3	1.2	0.6	0.7	0.1
Turkey	2.5	2.5	1.2	1.5	0.7	1.7	1.9	1.5	1.1	0.8	0.5	0.1
Uruguay	2.0	2.3	1.7	1.0	0.5	2.1	2.1	0.9	0.6	0.5	0.4	0.1
Venezuela	2.3	2.4	1.9	1.6	0.3	1.3	1.3	1.1	1.2	0.3	1.1	0.1

Table G.26: Robustness R6

	AUT	BEL	CHE	DNK	ESP	FIN	FRA	GBR	ITA	LUX	NLD	SWE
Australia			7.3			5.9					3.7	
Bahrain	8.0			2.5	10.4	7.0	3.2	11.5	7.7	23.8		9.5
Cameroon		4.7										
HongKongSAR	0.7				79.8			33.9	31.4	29.9		17.1
Japan	83.5	87.1	78.5	97.5	9.8	82.6	75.7	54.6	60.9	46.2	86.0	73.4
NewZealand	7.7	6.1					18.7					
SaudiArabia			14.3								10.4	
Zambia		2.1				4.5	2.5					

Table G.27: Robustness R7

	AUT	BEL	CHE	DNK	ESP	FIN	FRA	GBR	ITA	LUX	NLD	SWE
Cyprus					75.8			33.8	18.0	87.4		13.3
Ireland	100.0	100.0	100.0	100.0	24.2	100.0	100.0	66.2	82.0	12.6	100.0	86.7

Table G.28: Robustness R8

	AUT	BEL	CHE	DNK	ESP	FIN	FRA	GBR	ITA	LUX	NLD	SWE
Australia		4.3	6.5			19.5	7.6				3.5	
Bahrain	7.7			3.2	10.1	3.7	1.0	12.1	7.8	26.4		10.2
CotedIvoire		3.5										
HongKongSAR					78.9			34.7	31.6	33.6		17.7
Japan	82.1	78.1	78.4	96.8	8.3	65.2	64.2	53.2	60.5	39.9	85.9	72.2
NewZealand	10.1	14.1			2.6	11.7	27.1					
SaudiArabia			15.1								10.5	

Table G.29: Robustness R9

	AUT	BEL	CHE	DNK	ESP	FIN	FRA	GBR	ITA	LUX	NLD	SWE
Australia		1.1	8.2			15.3					3.7	
Bahrain	7.9	0.4		3.0	10.4	5.8	4.6	11.8	7.8	24.9		9.9
HongKongSAR	0.4				79.7			33.2	31.3	25.1		16.0
Japan	83.0	89.9	78.7	97.0	9.9	75.8	83.4	55.0	60.9	50.0	86.0	74.1
NewZealand	8.6						5.5					
Paraguay		8.7				3.2	6.5					
SaudiArabia			13.1								10.3	

Table G.30: Robustness R10

	AUT	BEL	CHE	DNK	ESP	FIN	FRA	GBR	ITA	LUX	NLD	SWE
Australia	11.5		25.0	4.7		25.8	9.6	21.6		16.9	4.2	1.9
Bahrain					8.4					20.9		
HongKongSAR					75.2			45.9	11.4	62.2		
NewZealand	88.5	95.1	63.9	95.3	16.4	74.2	89.7	32.4	88.6		83.5	98.1
SaudiArabia		4.9	11.1				0.6				12.4	