WORKING PAPERS IN ECONOMICS

No. 10/07

ARILD AAKVIK AND SIGVE TJØTTA

DO COLLECTIVE ACTIONS CLEAR COMMON AIR? THE EFFECT OF INTERNATIONAL ENVIRONMENTAL PROTOCOLS ON SULPHUR EMISSIONS

Department of Economics
UNIVERSITY OF BERGEN

This version: 13/07/2007

Do Collective Actions Clear Common Air?
The Effect of International Environmental Protocols on Sulphur Emissions

Arild Aakvik and Sigve Tjøtta[1]
Department of Economics, University of Bergen, Bergen, NORWAY

ABSTRACT
We consider the effects of voluntary international environmental protocols on emissions with regard to the 1985 Helsinki Protocol and the 1994 Oslo Protocol on the reduction of sulphur oxides. Our analysis utilizes panel data from 30 European countries for the period 1960–2002. We divide these countries into "participants" and "non-participants", i.e., those that did and those that did not ratify the specific protocol. We use a difference-in-difference estimator that focuses on the difference in emissions before and after signing a specific protocol and compares it with this difference for non-participant countries. Difference-in-difference estimation methods rely on annual data and may induce serial correlations in the variables. We use randomly generated placebo protocols to test the estimated effects. In a panel data regression model, where we include country and year dummies, the effect of the Helsinki agreement in reducing sulphur emissions is around three per cent per year, and the effect of the Oslo agreement is around four per cent per year. Correcting the standard errors for serial correlation in both dependent and independent variables is important and overlooked in the previous empirical literature on the evaluation of international agreements.

JEL Classifications: Q28, H49, D62

Keywords: International agreements, programme evaluation, placebo, serial correlation

---

[1] Contact information: Sigve Tjøtta, Department of Economics, Herman Foss gt. 6, 5007 Bergen, Norway. Email: sigve.tjotta@econ.uib.no. Telephone: +47 55 58 92 20.

1. Introduction

Success in managing global public goods and commons is important for our future welfare. Examples of global public goods include global warming, maintenance of international macroeconomic stability, international trade rules, international political stability, humanitarian assistance, and knowledge. Institutions for managing international public goods include international environmental agreements such as the Kyoto Protocol on emission reduction of greenhouse gases, the World Bank, the International Monetary Fund, the World Trade Organization, and the United Nations. Public goods crossing national jurisdictional borders add a dimension to Samuelson's (1954) general theory for public goods. Under the current international law, obligations may be imposed only on a sovereign state with its consent. Hence, multinational institutions and international agreements often have weak or even lack explicit control and enforcement mechanisms. Compliance with agreements is often hard to control and verify, and moreover, there are seldom explicit sanction mechanisms in these agreements. With this in mind, it is reasonable to address the question of whether these institutions work or not.

In this paper, we address this question by evaluating two specific international environmental agreements: The 1985 Helsinki Protocol on the Reduction of Sulphur Emissions or their Transboundary Fluxes by at least 30 per cent, and The 1994 Oslo Protocol on Further Reduction of Sulphur Emissions. In the 1985 Helsinki Protocol, each signatory should reduce sulphur dioxide emissions by at least 30 per cent compared with the 1980 level "as soon as possible and at the latest by 1993" according to Article 2 in the protocol.[2] The 1994 Oslo Protocol is a direct follow-up protocol of the 1985 Helsinki Protocol, but participation is not restricted to countries participating in the first sulphur protocol. The 1994 Oslo Protocol specifies different emission ceilings expressed as annual emission ceilings for the years 2000, 2005, and 2010. As both protocols are voluntary in compliance, they share many of the same characteristics as most institutions managing global public goods and commons.

Even though there are contradictory results from studies evaluating international environmental agreements, most of the economics literature concludes that they "tend to codify Nash behaviour and, as such, do not present much of a co-operative gain" (Arce *et al.* 2001, p. 494).[3] As far as we know, the 1994 Oslo Protocol has not yet been evaluated by an econometric approach, but the 1985 Helsinki Protocol has undergone quantitative assessment. Levy (1993), Murdoch *et al*. (1997) and Helm and Sprinz (2000) find that the 1985 Helsinki Protocol reduced sulphur emissions further than expected reductions without the protocol. Ringquist and Kostadinova (2005) criticizes these studies for failure to take into account the fact that ratification of the Protocol is not random but a result of a self-selection decision process. They find that the 1985 Helsinki Protocol has had no significant effect on sulphur emissions.

To control for the self-selection process, we use programme evaluation methods, and difference-in-difference estimations in particular, to control for observed and unobserved

---

[2] For the protocol's webpage, see http://www.unece.org/env/lrtap/status/lrtap_s.htm

[3] These arguments are often based on acceptance of the hypothesis that countries act according to the non-cooperative Nash model by theoretical/simulation methods (Mähler (1990), Hoel (1992), Barrett (1994, 2001, 2003), Carraro and Siniscalco (1993, 1998), and Finus and Tjøtta (2003)). In contrast, Lange and Vogt (2003) explain the large number of observed international cooperations by arguing that these are driven by preference for equity in addition to self-interest. Econometrical approaches (Murdoch and Sandler (1997), Murdoch *et al.* (1997), Bratberg *et al.* (2005)) find contradicting results.

factors affecting emission levels.[4] Difference-in-difference estimation focuses on the difference in emissions before and after signing a specific protocol and compares it with this difference for non-signatory countries. Identification of the effect is obtained by repeated annual observations of emissions from signatories and non-signatories as a control group both before and after the signing date. We use a very flexible panel data model that allows for country-specific growth effects, and standard country and year effects.

We expand on the econometric framework of Ringquist and Kostadinova (2005) by introducing both the 1985 Helsinki Protocol and the 1994 Oslo Protocol for reductions in sulphur emissions. In addition, we use a larger time span for analysis. Most importantly, we look at relative changes in emissions rather than absolute levels in the difference-in-difference model. This is important in models with potential selection bias because we allow for both the level and yearly changes in emissions to be different for the different treatment groups prior to the signing period. In addition, difference-in-difference estimation methods rely on annual data and this may induce serial correlations in the variables. In particular, the variable in evaluation focus—signing an agreement—will be serial correlated. Overlooking this may, as demonstrated by Bertrand *et al.* (2003), introduce biased estimates of the standard errors. To overcome this bias, we follow Bertrand *et al*. (2003) and use for inference the empirical distributions of estimated effects for placebo signatories.

---

[4] See, e.g., Heckman and Robb (1985), Heckman and Hotz (1989), or Moffitt (1991) for an overview of econometric evaluation methods for panel data.

Our results contradict those found in Ringquist and Kostadinova (2005) for the 1985 Helsinki Protocol. In a panel data regression model, where we include country and year dummies, and look at relative changes in emissions within the difference-in-difference framework, the effect of the Helsinki agreement in reducing sulphur emissions is estimated to be around three per cent per year, and the effect of the Oslo agreement is estimated to be around four per cent per year. These effects are significantly different from zero in most of our model specifications.

The rest of the paper is organized as follows. In section 2, we explain the historical background for the two protocols. In section 3, we outline our econometric approach, emphasizing programme evaluation principles and methods for utilizing placebo protocols. In section 4, we present data, and in Section 5, we discuss the results. Finally, Section 6 brings some concluding remarks.

2. Background

Environmental problems that cross national borders require some form of international cooperation. The large number of multilateral agreements on environmental problems is evidence that this challenge is addressed. By 1994, according to Barrett (1994, p. 878), more than 100 multinational environmental agreements were in force. This includes agreements on marine fisheries and pollution, international rivers, lakes and groundwater, conservation of species, and protection of pet animals. Some of these agreements are regional, such as the *Agreement Concerning the International Commission for the Protection of the Rhine Against Pollution*, which came into force in 1965 with five countries. Others are global agreements, such as the *United Nations Framework*

Convention on Climate Change adopted in 1992, where the number of signatories is 189 countries.

Most multinational environmental problems are characterized by conflict of interest and/or coordination problems. There is conflict of interest as each country may have an incentive to free-ride on emission reductions from other countries. As a consequence, countries that would benefit from cooperation end in a prisoners' dilemma situation. Even without conflict of interest, environmental and management problems of other commons may be difficult to overcome as there may be problems with coordination. A government may find it hard to contribute without credible contributions from other governments.

The conflict of interest and control problems may be partly solved as governments engage in repeated intergovernmental relations in managing environmental commons and other relations, such as EU integration, NATO enlargement, and international trade agreements. These relations may overcome incentives to free-ride. Ostrom (1990) argues, using numerous examples, that local communities handle incentives to free-ride by voluntary cooperation when supplying public goods. In the same manner, governments may manage to deal (partly) with incentives to free-ride.

Success or failure of international agreements depends on participation and compliance incentives. A specific agreement can fail to incorporate all relevant countries that should have been part of the agreement. Depending on the group of participating countries, the agreement can also fail to implement efficient emission policies.

The 1979 Geneva Convention and its follow-up protocols

For many European countries, most deposits of sulphur and nitrogen oxides are airborne from other countries. These air pollutants can be carried by winds for hundreds, even thousands, of kilometres before being deposited in soil or water. These pollutants cause damaging effects on fish stocks and forests, and cause health threats to people by reducing the air quality. Clearly, nitrogen and sulphur oxide emissions in one country create environmental and health consequences in other countries.

To deal with these problems of air pollution, a group of European and North American countries agreed upon *The 1979 Geneva Convention on Long-range Transboundary Air Pollution.* When marking its 25th anniversary in 2004, the number of parties to the agreement was 49 including the European Union. The 1979 Geneva Convention was extended by eight follow-up protocols[5] that specify specific targets for abatements for the parties to the protocols. The first of these was *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*. This protocol has three main components: collection of emission data for sulphur and nitrogen oxide and other air pollutants, measurement of air and precipitation quality, and modelling of atmospheric dispersion.[6] The Protocol also requires mandatory contributions supplemented by voluntary contributions, for sharing the cost of financing the monitoring of emissions and the quality of specific air pollutants. The mandatory contribution is given as a percentage

---

[5] For the 1979 Geneva Convention and its follow-up protocols, see http://www.unece.org/env/lrtap/welcome.html and all eight follow-up protocols.
[6] According to the official protocol webpage, see http://www.unece.org/env/lrtap/emep_h1.htm.

of total cost. The protocol came into force on 28 January 1988, and by June 2004, it encompassed 41 parties.

There are three follow-up protocols regulating sulphur emissions: The 1985 Helsinki Protocol on the Reduction of Sulphur Emissions or their Transboundary Fluxes by at least 30 per cent, The 1994 Oslo Protocol on Further Reduction of Sulphur Emissions, and The 1999 Gothenburg Protocol to Abate Acidification, Eutrophication and Groundlevel Ozone. In the 1985 Helsinki Protocol, each signatory should reduce sulphur dioxide emission by at least 30 per cent compared with the 1980 level "as soon as possible and at the latest by 1993" according to Article 2 in the protocol.[7] It came into force on 2 September 1987 and has 22 parties. The 1994 Oslo Protocol is a direct follow-up protocol of the 1985 Helsinki Protocol, but participation is not restricted to countries participating in the first sulphur protocol. For example, Greece, Ireland, Poland, and the United Kingdom did not sign the 1985 Helsinki Protocol but did sign the 1994 Oslo Protocol. The 1994 Oslo Protocol specifies different emission ceilings expressed as annual emission ceilings for the years 2000, 2005 and 2010. It became effective on 5 August 1995 with 25 parties. The 1999 Gothenburg Protocol sets emission ceilings on sulphur, nitrous oxides (NO_X) and volatile organic compounds (VOC) as the other major air pollutants responsible for the formation of ground-level ozone and ammonia. It is, however, not yet in effect.

Among the follow-up protocols, there are two specific protocols on NO_X emissions: *The 1988 Sofia Protocol concerning the Control of Nitrogen Oxides or their Transboundary*

---

[7] For the protocol's webpage, see http://www.unece.org/env/lrtap/status/lrtap_s.htm

Fluxes, and the aforementioned 1999 Gothenburg Protocol. Finally, there are three protocols concerning other pollutants: *The 1991 Geneva Protocol concerning the Control of Emission of Volatile Organic Compounds or their Transboundary Fluxes, The 1998 Aarhus Protocol on Heavy Metals,* and *The 1998 Aarhus Protocol on Persistent Organic Pollutants (POPs).*

3. Empirical modelling

Our aim is to evaluate the effect of a specific agreement on the signatories' sulphur emissions. In particularly, we investigate whether signatories reduce their emissions more than they would have done without this agreement. Clearly, reduction of emissions is a restricted measurement of success, and many agreements have other goals. For example, the 1994 Oslo Protocol on sulphur emissions also focuses on damage caused by the pollutants, as the deposition of oxidized sulphur should not exceed specified critical loads (see Appendix 2), and cost-efficient emission reduction. As we are focusing on emission reduction, we overlook success along other dimensions of a protocol. However, we believe that reduction in emissions is one of the key measurements of success.

We measure the effect over a long time span, as the intention of these protocols is to change emission policy over a long period. In the 1985 Helsinki Protocol, the aim is to reduce sulphur dioxide emissions by at least 30 per cent as soon as possible but at the latest by 1993. Similar, the 1994 Oslo Protocol sets up specific reductions by the years 2000, 2005 and 2010.

We model the emission process as a random growth model (see, for example Heckman and Hotz, 1989), and to formalize we define for year t = 0, 1, 2,…, T and country i = 1, 2,…, N the following variables.

Y_{it} = a country i's natural logarithm of sulphur emissions in year t

X_{it} = a vector of covariates such as GDP and population

T_p = a set of years where protocol p may have an effect, where $\mathrm{T_p} \subseteq \{0, 1, 2, ..., \mathrm{T}\}$

τ_{p} = the first year when protocol p may have an effect

$$D_{it}^{p} = \begin{cases} 1 \text{ if country } i \text{ signed protocol } p \text{ and } t \in T_p \\ 0 \text{ otherwise} \end{cases}$$

The random growth model for county i at time t is given by:

$$Y_{it} = c_i + g_i t + \delta^1 D_{it}^1 [t - (\tau_1 - 1)] + \delta^2 D_{it}^2 [t - (\tau_2 - 1)] + X_{it}\beta + v_{it} \text{ for } t = 1, 2, \ldots, T, \quad (1)$$

where the parameters c_i and g_i are country fixed effects and country annual growth effects, respectively. The *annual* effects of the two protocols are captured by the parameters (δ^1, δ^2). Differentiating the random growth model in (1) with respect to t, we obtain:

$$\Delta Y_{it} = g_i + \delta^1 D_{it}^1 + \delta^2 D_{it}^2 + \Delta X_{it}\beta + \Delta v_{it}, \quad (2)$$

where Δ is the time difference operator, for example $\Delta Y_{it} = Y_{it} - Y_{it-1}$. The interpretation of the coefficients in the level model (1) and the difference model (2) is the same; however, the assumption about the error term is different. This will have consequences for the estimation of the standard error of the coefficients in our model. In particular, model (1) is generally ridden with serial correlation (see, for example, Wooldridge, 2002; Bertrand *et al*., 2003; Helland and Tabarrok, 2004). Both specifications are estimated using OLS with country-specific fixed effects.

The treatment effect of protocol *p* at time $t \in T_p$, i.e., δ^1 and δ^2, is interpreted as an annual effect. The identifying assumptions for δ^1 and δ^2 to be causal effects are as follows. (i) The change over time in sulphur emissions in the contrafactual situation is the same as it is for the control group, i.e., $\mathrm{E}(\Delta_t Y_{it}^c \mid D_{it}^p = 1) = \mathrm{E}(\Delta_t Y_{it} \mid D_{it}^p = 0)$. (ii) Pre-treatment relative change for the group of signatories is the same as it would have been without signing the protocol, i.e., there is no adjustment in the pre-signing period. Hence, as an estimator for the annual effect for protocol *p* we utilize:

$$\delta^p = \mathrm{E}(\Delta_t Y_{it} \mid D_{it}^p = 1) - \mathrm{E}(\Delta_t Y_{it} \mid D_{it}^p = 0),$$

which is the difference-in-difference model for yearly changes in emissions.

Serial correlation and randomized protocols

As we remarked in the introduction, estimation of equation (1) by OLS is subject to a possible serial correlation. Clearly, as we model emissions using time differences (ΔY_{it}), we reduce the potential serial correlation in the dependent variable compared with estimating a model using the *level* of emissions (Y_{it}) as the dependent variable. However, the treatment variables in themselves, D_{it}^p, are highly correlated over time for a given country. If country i signed protocol p in year t, it will also have signed the protocol in the years afterwards. This serial correlation will induce biased estimates of the standard errors of the estimated effects ($\hat{\delta}^p$) (see Wooldridge, 2002; Bertrand *et al.*, 2003). The direction of the bias depends on serial correlations in the error terms and the explanatory variables. In a simple model with one explanatory variable with a first-order serial correlation in the independent variable and a first-order correlation for the error term, the bias is downward if the serial correlation in the error term and independent variable are positive. The estimated standard errors are smaller than true standard errors in this case. If the first-order correlation in error terms is negative, the estimated standard errors are too large.[8]

To overcome the potential problem with serial correlation, we follow Bertrand *et al.* (2003) and use randomization inference by estimating the effect and standard deviation of placebo signatories. We expand the placebo framework used in the literature and suggest several different randomization strategies. In general, we generate placebo protocols P by randomly assigning a protocol by country and year, P_{it}^{p}, and estimate the effects using OLS.

$$\Delta Y_{it} = g_i + \gamma^1 P_{it}^1 + \gamma^2 P_{it}^2 + \Delta X'_{it}\beta + \Delta_t \varepsilon_{it} \tag{3}$$

We generate 1000 random placebo protocols, and generate the distributions of γ^p denoted by $G^p(.)$ for p = 1, 2. Assuming the model is correctly specified, the mean of the

[8] Nicholls and Pagan (1977) analyse the problem using alternative assumptions about the serial correlation for the independent variable and the error term.

distribution is zero because randomly choosing protocols should have no effect on emissions. The average number of randomly drawn country-years is the same as the number in the actual sample signing the 1985 Helsinki Protocol and the 1994 Oslo Protocol.[9] We perform 1000 independent draws of (P_{it}^1, P_{it}^2). For each draw, we run the regression in equation (3).

The randomization of placebo protocols is drawn in four different ways depending on assumptions about both the data-generating process and the structural properties of the parameters in the model. (i) We assume that the effect of a protocol, i.e., the parameter γ_p, is *independent* of time. This assumption means that the effect of the protocol in the first year is the same whether the protocol started in 1970 or in 1990.[10] We label this assumption that the effect is structural or path independent. In the randomization algorithm, a placebo protocol is drawn randomly for each country-year in the whole observation period 1961–2002 such as the average country-years active placebo protocols are the same as in the actual sample. For example, the first drawn placebo protocol may be the year 1961 for a specific country but not the following year for the same country. (ii) We still assume that the effect is structural. However, we restrict the placebo protocols to be in place in the same consecutive number of years as the actual protocol. Each year we randomly draw a country to be a part of a placebo protocol or not. If a country is drawn to be part of the placebo protocol, it will be part of the protocol for succeeding years such that the total average number of years is the same as that for the

---

[9] In our data there are in total 1,038 country-year observations. Out of these, there are 123 country-years where the 1985 Helsinki Protocol's indicator $D_{it} = 1$, and 158 country-years for the 1994 Oslo Protocol.
[10] In addition, we assume that the effect is the same from the first to the second year and from the second to the third year, etc.; i.e., constant over time.

observed 1985 Helsinki Protocol and 1994 Oslo Protocol. (iii) We assume that the effect is *restricted* to the actual period of the protocols, i.e., the parameter γ_p is only defined in the actual protocol period, T_p. This assumption, we take to mean, as the effect is path dependent and not a structural phenomenon. In the randomization algorithm, a placebo protocol is drawn randomly for each $t \in T_p$. This implies that a country may be in a placebo protocol p in the first year T_p but may not be in the second year. (iv) Finally, we assume that the effect is restricted to the actual signatory period as in (iii), and that the effect of protocol *p* is defined for the *whole* period T_p. A placebo protocol is drawn randomly for the whole period T_p. This means that if a country is drawn to be in a placebo protocol the first year in the actual protocol period T_p, the protocol indicator variable, P_{it}, is constant for the rest of the actual period T_p.

Even though we report standard errors for all four randomization strategies, they differ in the assumption about the data-generation process and whether our parameters are assumed to be structural or not. Randomization strategy (iv) replicates our data perfectly except that the protocol country is randomly drawn and the placebo effect is restricted to an actual period. However, the time dimension is irrelevant if the effect of a protocol is a structural parameter, meaning that the effect is not time dependent and we allow for it to be in other periods as well. In this case, we draw placebo protocols for the entire observation period (1960–2002) as in randomization strategies (i) and (ii). In (ii), we assume that the protocol period is fixed, which is not the case in (i).

To test the null hypothesis of no effect of signing a protocol, we check whether the estimate of δ^p from equation (2) lies in the distribution $G^p(.)$. For example, to form a two-level test at the five per cent level, we find γ^p at 2.5 per cent of the lower and upper tail of the distribution. If our estimate of δ^p is inside the interval defined by the two points, we cannot reject the null hypothesis of no-effect. If it lies outside this interval, we reject the null hypothesis of no-effect. As we do not expect the effect of protocols to be counterproductive, i.e., the estimate of δ^p to be positive (increased emissions), we also use a one-sided test. Formally, we define the null hypothesis as $\delta^p \geq 0$ with alternative hypothesis $\delta^p < 0$.

4. Data

Our analysis of the sulphur protocols utilizes data from 30 European countries in the period 1960–2002. All countries in the data set have signed and ratified the 1970 Geneva Convention. We include a country as a "signatory" of a protocol if it had *both* signed and ratified the particular protocol. The ratification has to be in 1993 for The 1985 Helsinki Protocol and in 2002 for the 1994 Oslo Protocol. Of these 30 countries, 18 signed the 1985 Helsinki Protocol, 19 signed the 1994 Oslo Protocol, 13 signed both protocols, and six signed neither of these protocols (see Table 1). The 1985 Helsinki Protocol was signed on 9 July in 1985, and we assume that the period of potential effect is 1986–93. The 1994 Oslo Protocol was signed on 14 June in 1994, and we assume that the effect is from 1995 until 2002.

(Table 1 about here)

Data on sulphur emissions are based on Stern (2006), which is again based on EMEP's webpage.[11] For some additional data, we also use EMEP's webpage. The population and gross domestic products (GDP) data are taken from PENN World Table version 6.2, Heston *et al.* (2006). The political changes in Europe might have had an effect on emissions behaviour. Therefore, we include a control dummy variable for the former communist countries of Eastern Europe, reflecting the transition process. We use 1991 as the first transition year, indicating post-transition with a dummy. For a more detailed description of the data, see Appendix A.

(Figures 1 and 2 about here)

Figures 1 and 2 show sulphur emissions per capita, GDP per capita, and sulphur emissions per GDP dollar for the signatories of the 1985 Helsinki Protocol and the group of control countries in the period 1960–2002 (Figure 1), and the signatories of the 1995 Oslo Protocol and the control group (Figure 2). The GDP per capita is increasing for both signatories and the control group, but the level of GDP per capita is lower for the controls. The control groups use sulphur more intensively in their economies as sulphur emissions per GDP are higher than the signatories of the two protocols. Sulphur emissions increased for all countries from 1960 until at least 1975. From the beginning of the 1980s, when discussions of the 1985 Helsinki Protocol began, most countries had already started to reduce the level of sulphur emissions. This decline is also confirmed by Figure 3, which shows median relative emission changes over time for the two protocols.

---

[11] The Co-operative Programme for Monitoring and Evaluation of the Long-Range Transmission of Air Pollutants in Europe (the EMEP programme) monitors emissions regulated by the 1979 Geneva Convention. The data are from http://www.emep.int/.

For both protocols, the annual growth for the median signatory country is below that for the median non-signatory countries in the protocol periods, but even though the picture for pre-protocols is somewhat mixed, it seems as though annual growth for median signatory countries is below that for the median non-signatory countries. Summing up, the raw data seem to indicate that both protocols worked. For more accurate statements, we now turn to the econometric analyses.

(Figure 3 about here)

5. Results and analyses

The results from the regression model given in differentiated form in equation (2) are given in Table 2 below. Except for the participation and transition country dummies, the variables are in logarithm and time differenced; thus, their coefficients are interpreted as elasticities. Our preferred specification of the model includes controls for both year effects and country-specific growth effects in the differentiated model. We also include the results from a regression with no year and country growth effects (model I in Table 2), and a regression model where we include year effects but not country-specific growth effects (model II in Table 2). It turns out that allowing for country-specific growth effects in emissions is important, because countries vary substantially in emission trajectories. Thus, in the discussion, we will focus on model III in Table 2, because this offers the most flexible modelling of changes in emissions over time.

From Table 2, model III, we see that a one per cent increase in gross domestic product will increase emissions by 0.45 per cent. The population variable is not significantly different from zero in our preferred model (model III). In the former communist countries in Eastern Europe, the annual relative change in emissions is between minus three and four per cent annually. However, this effect is not statistically different from zero in model III where we allow for country-specific growth effects.

(Table 2 about here)

Turning to the main variables of this paper, we see that, without controlling for time and country effects (model I in Table 2), the estimated annual effect of the 1985 Helsinki Protocol in the period 1986–1994 is –0.064, corresponding to a 6.4 per cent annual reduction in emissions. The estimated effect of the 1994 Oslo Protocol is an 8.4 per cent annual reduction in emissions. The estimated robust standard errors are 0.011 for the Helsinki and the Oslo estimates, and thus both are highly significant.

Model III in Table 2 includes both a common yearly growth rate and a country-specific yearly growth rate in emissions independent of the protocol variables, and is our preferred model. We see that the effect of the protocols is reduced to 3.1 per cent for the Helsinki Protocol and 4.0 for the Oslo Protocol. The robust standard errors for these estimates are 0.021 and 0.025 respectively. Thus, the coefficient estimates are not significant at the 10 per cent level for a two-sided test.

To investigate the serial correlation for the dependent variable, we estimate serial correlations from regressions on time differences of emissions ($\Delta_t Y_{it}$) controlling for covariates ($\Delta_t X_{it}$) from the model given by equation (2). The serial correlation coefficients are obtained by a regression of the residuals on the corresponding lagged residuals. The estimated serial correlation coefficients are small in the model where we look at yearly changes in emissions within the difference-in-difference model (see Appendix B). The first-order serial correlation coefficients are around –0.03 in Table B2, and they are not significantly different from zero. The correlation coefficients between the residual and lagged values of the residuals are also very small (see Table B.3).

The serial correlation coefficients in Table B.2 are considerably smaller than similar correlation coefficients obtained by regression on the corresponding *level* model usually used in the literature, i.e., regression of emission (Y_{it}) controlling for covariates $(\mathrm{X}_{\mathrm{it}})$, and year and country dummies (see Table C.2 in Appendix C). This model is estimated from equation (1) above. The estimated first-order coefficients are around 0.8 and strongly significant. All this demonstrates that serial correlations related to the dependent variable are reduced when we use time differences in logarithmic emissions and covariates as variables instead of level of emissions and level of covariates. The standard error from the level model is much smaller than the standard error for the difference model. This is according to theory. Positive serial correlation in the error term will cause standard errors to be understated, and this will be exaggerated if we also have serial correlation in the independent variables (see Helland and Tabarrok, 2004). The problem with serial correlation is much smaller for the difference model given in equation (2) compared to the level model in equation (1).

We estimate the standard errors for the Helsinki and Oslo Protocols in case of serial correlation in the dependent and/or independent variables using placebo protocols. The empirical distributions of the effect of the placebo protocols are given in Figure 4 for the

Helsinki Protocol and in Figure 5 for the Oslo Protocol. As expected, the placebo effects are normally distributed with mean value zero. The standard errors for the four different randomization strategies are given in Table 3 for the Helsinki Protocol and the Oslo Protocol.

The standard errors vary slightly depending on our randomization strategy. The standard error increases monotonically from specification (i) to specification (iv) (see Table 3 and Figures 4 and 5). In addition, we find that the standard errors are smaller in the placebo experiments compared with the robust standard errors as reported in Table 2, except for the standard error for the Oslo Protocol using randomization strategy (iv).

The estimated standard errors for the Helsinki Protocol and the Oslo Protocol using randomization strategy (i) are 0.013 and 0.011, respectively. Thus, the estimated yearly effects of both protocols are significantly different from zero at the one per cent level. If we use randomization strategy (ii), the estimated standard errors are 0.016 and 0.014 for both the Helsinki and Oslo Protocols, respectively. Thus, the Helsinki Protocol is significant at the 10 per cent level, while the Oslo Protocol is significant at the five per cent level. The same conclusion is reached using randomization strategy (iii), where the estimated standard error for both protocols is 0.018. The estimated effect for the Helsinki Protocol using randomization strategies (ii) and (iii) is significant at the five per cent level for a one-sided test, while the Oslo Protocol is significant at the one per cent level for a one-sided test. Lastly, the estimated standard errors from strategy (iv) are 0.019 and 0.031 for the Helsinki and Oslo Protocols, respectively. Thus, the estimated coefficients ($\hat{\delta}^1, \hat{\delta}^2$) are not significantly different from zero in this case.

As we mentioned earlier, we do not expect the effect of protocols to be counterproductive, i.e., the estimate of δ^p to be positive (increased emissions). Thus, we also use a one-sided test. Formally, we define the null hypothesis as $\delta^p \geq 0$ with alternative hypothesis $\delta^p < 0$. However, in a one-tailed test (not reported here), both protocols are significant at the 10 per cent level for all six standard errors.

(Figures 4 and 5 about here)

(Table 3 about here)

6. Concluding remarks

Our opening question was whether voluntary multinational environmental agreements succeed in reaching their aims. Based on our analysis of the 1985 Helsinki Protocol and the 1995 Oslo Protocol on sulphur emissions, our answer to this question is “yes” conditional on fundamental assumptions in our analysis. Our analysis utilizes panel data from 30 European countries for the period 1960–2002. We divide these countries into “participants” and “non-participants”, i.e., those that did and those that did not ratify the specific protocol. We use a difference-in-difference estimator that focuses on differences in emissions before and after signing a specific protocol and comparing this with the differences for non-participating countries. We find that the annual reduction in emissions due to the Helsinki Protocol is three per cent, and four per cent for the 1995 Oslo Protocol. To overcome the problem of biased standard errors, we follow Bertrand *et al*. (2003) and compute the difference-in-difference estimates for 1000 randomly generated placebo protocols. We include several different randomization strategies depending on the data-generating process and assumptions about the parameters in our model. We use the empirical distribution of placebo effects to test whether the estimated effects are significantly different from zero. The estimated effects are significant in some of our models, but not robust to different randomization strategies for placebo protocols. Thus, even though the annual effects of the two protocols in terms of reduced sulphur emissions are estimated at three per cent for the Helsinki Protocol and four per cent for the Oslo Protocol, the estimated standard errors indicate that a definitive conclusion about the effectiveness of these protocols cannot be reached.

References

Arce, M., Daniel, G., Sandler, T., 2001. Transnational public goods: strategies and institutions. European Journal of Political Economy 17, 493–516.

Barrett, S., 1994. Self-enforcing international environmental agreements. Oxford Economic Papers 46, 878–894.

Barrett, S., 2001. International cooperation for sale. European Economic Review 45, 1835–1850.

Barrett, S., 2003. Environment and Statecraft: The Strategy of Environmental Treaty-Making. Oxford University Press, Oxford.

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

Bertrand M., Duflo, E., Mullainathan, S., 2003. How much should we trust differences-in-difference estimates? Quarterly Journal of Economics 199, 249–275.

Carraro, C., Siniscalco D., 1993. Strategies for the international protection of the environment. Journal of Public Economics 52, 309–328.

Carraro, C., Siniscalco, D., 1998. International environmental agreements: incentive and political economy. European Economic Review 42, 561–572.

EMEP, 1997. MSC-W Status Report 1/97 Part II, Meteorological Synthesizing Center-West, The Norwegian Meteorological Institute, Oslo, Norway.

Finus, M., Tjøtta, S., 2003. The Oslo Agreement on sulfur reduction in Europe: The great leap forward? Journal of Public Economics 87, 2031–2048.

Heckman, J.J., Hotz, J.V., 1989. Choosing among alternative nonexperimental methods for estimating the impact of social programs: the case of manpower training. Journal of the American Statistical Association, 84 (408), 862–874.

Heckman, J.J., Robb, R., 1985. Alternative methods for evaluating the impact of interventions: An overview. Journal of Econometrics 30 (12), 239–267.

Helland, E., Tabarrok, A., 2004. Using placebo laws to test “more guns, less crime”: A note. Advances in Economic Analysis & Policy: The Berkeley Electronic Press Journals. 4 (1), Article 1.

Helm, C., Sprinz, D., 2000. Measuring the effectiveness of international environmental regimes. Journal of Conflict Resolution 44 (5), 630–652.

Heston A., Summers R., Bettina A., 2006. Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006.

Hoel, M., 1992. International environmental conventions: The case of uniform reductions of emissions. Environmental and Resource Economics 2, 141–159.

Lange, A., Vogt, C., 2003. Cooperation in international environmental negotiations due to a preference for equity. Journal of Public Economics 87, 2049–2067.

Levy, M.A., 1993. European acid rain: The power of tote-board diplomacy. In: Haas, P.M., Keohane, R., Levy M.A. (Eds.), Institutions for the Earth: Sources of Effective International Environmental Protection. Cambridge MIT Press, pp. 75–132.

Mähler, K.A., 1990. International environmental problems. Oxford Review of Economic Policy 6, 80–108.

Moffitt, R., 1991. Program evaluation with nonexperimental data. Evaluation Review 15, 291–314.

Murdoch, J.C., Sandler T., 1997. The voluntary provision of a pure public good: The case of reduced CFC emissions and the Montreal Protocol. Journal of Public Economics 63, 331–349.

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

Nicholls, D.F., Pagan A.R., 1977. Specification of the disturbance for efficient estimation—an extended analysis. Econometrica 45 (1), 211–217.

Ostrom, E., 1990. Political Economy of Institutions and Decisions. Cambridge, Cambridge University Press.

Ringquist, E.J., 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.

Samuelson, P.A., 1954. The pure theory of public expenditure. The Review of Economics and Statistics 36 (4), 387–389.

Wooldridge, J.M., 2002. Econometric Analysis of Cross Section and Panel Data, the MIT Press, Cambridge, MA.

Table 1. Participating countries in the protocols and controls.

Helsinki & Oslo	Helsinki	Oslo	None
Austria	Bulgaria	Czech Republic, 1991–02	Iceland
Belgium	Russia, 1990	Greece	Poland
Denmark	USSR, 1990	Ireland	Portugal
Finland	West Germany, 1960–89	Slovakia, 1991–02	Romania
France	East Germany, 1970–89	Spain	Turkey
Hungary		United Kingdom	
Italia			
Germany, 1990–2002			
Luxemburg			
Netherlands			
Norway			
Sweden			
Switzerland			

Table 2. Estimate of relative changes in emissions using panel data (robust standard error in parenthesis).

Dependent variable: Annual change in log emission			
	I	II	III
Helsinki 1986–1993	–0.062***	–0.062***	–0.031
	(0.010)	(0.018)	(0.021)
Oslo 1995–2001	–0.085***	–0.093***	–0.040
	(0.011)	(0.016)	(0.025)
Gross domestic product	0.624***	0.532***	0.446**
	(0.149)	(0.174)	(0.188)
Population	1.983**	1.012	–0.530
	(0.862)	(0.892)	(1.315)
Transition = 1 after 1990	–0.032**	–0.040**	–0.037
	(0.013)	(0.016)	(0.023)
Year effects	No	F = 3.47***	F = 3.43***
Country growth effects (g_i)	No	No	F = 3.04***
Constant	–0.030***	0.023	0.012
	(0.009)	(0.024)	(0.027)
R-square	0.134	0.227	0.269
N	1,038	1,038	1,038

Notes

(1) Dependent variable: First-order differences in log emission.

(2) All independent variables are in first-order difference in logarithms except dummy variables.

(3) Significant levels two-tailed test: * 10 per cent, ** 5 per cent, *** 1 per cent based on robust standard errors.

Table 3. Estimated effects and different standard errors (in percentages).

Protocols:	THE 1995 HELSINKI	THE 1994 OSLO
Actual protocol period	1986–1993	1995–2002
Estimated annual effect (%)	–3.1	–4.0
Estimated standard errors	2.0	2.5
Robust standard errors	2.1	2.5
Standard errors from distribution of effects of placebo protocols for different assumptions:		
Structural effect parameter:		
Annual draw	1.3***	1.1***
Period draw	1.6*	1.4**
Effect parameter restricted to actual protocol period:		
Annual draw	1.8*	1.8**
Period draw	1.9	3.1

Notes

(1) Significant levels two-tailed test: * 10 per cent, ** 5 per cent, *** 1 per cent.

(2) Period draw means that a country is randomly drawn to participate in a placebo protocol for a period with length as actual protocol period.

Figure 1. Mean values for sulphur emissions per capita, GDP (1996 US dollars), and sulphur emissions per GDP for signatories of the 1985 Helsinki Protocol and control group.

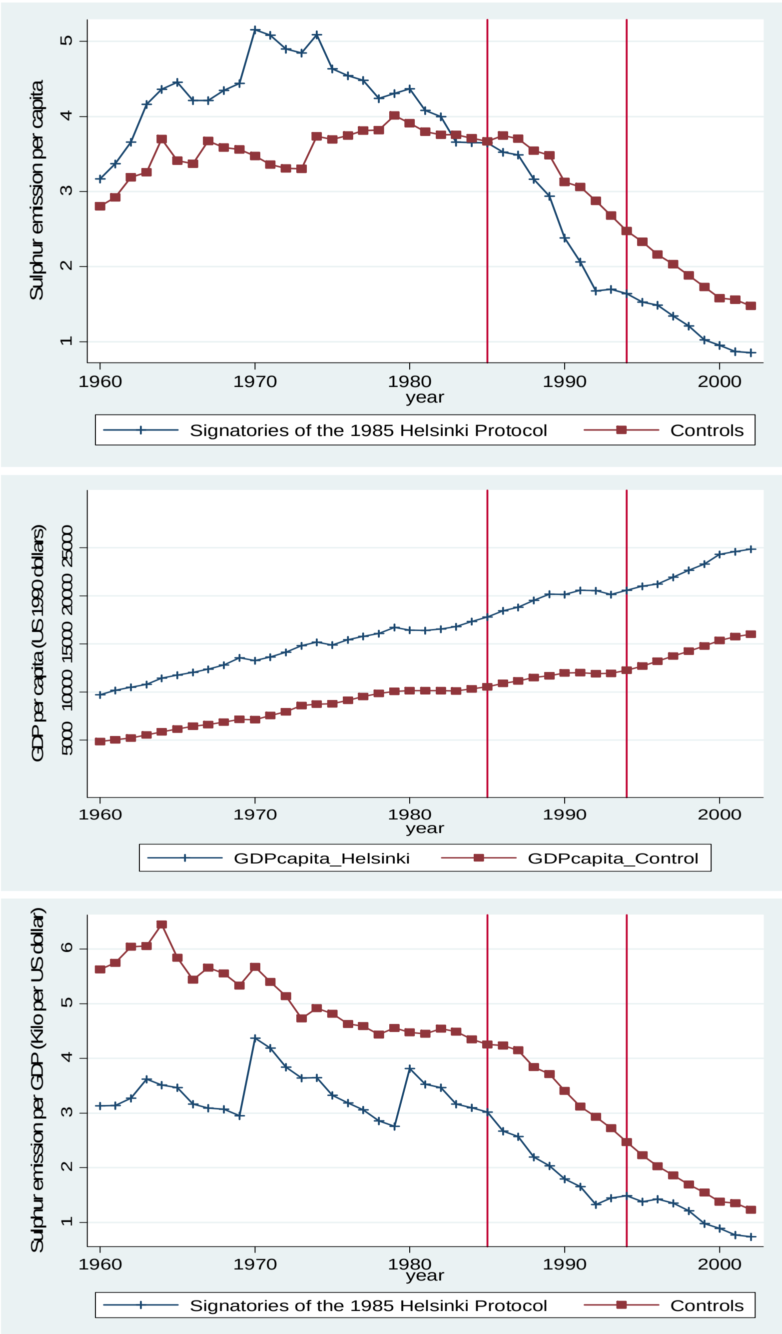


Figure 2. Mean values for sulphur emissions per capita, GDP per capita, and sulphur emissions per GDP for signatories of the 1994 Oslo Protocol and control group.

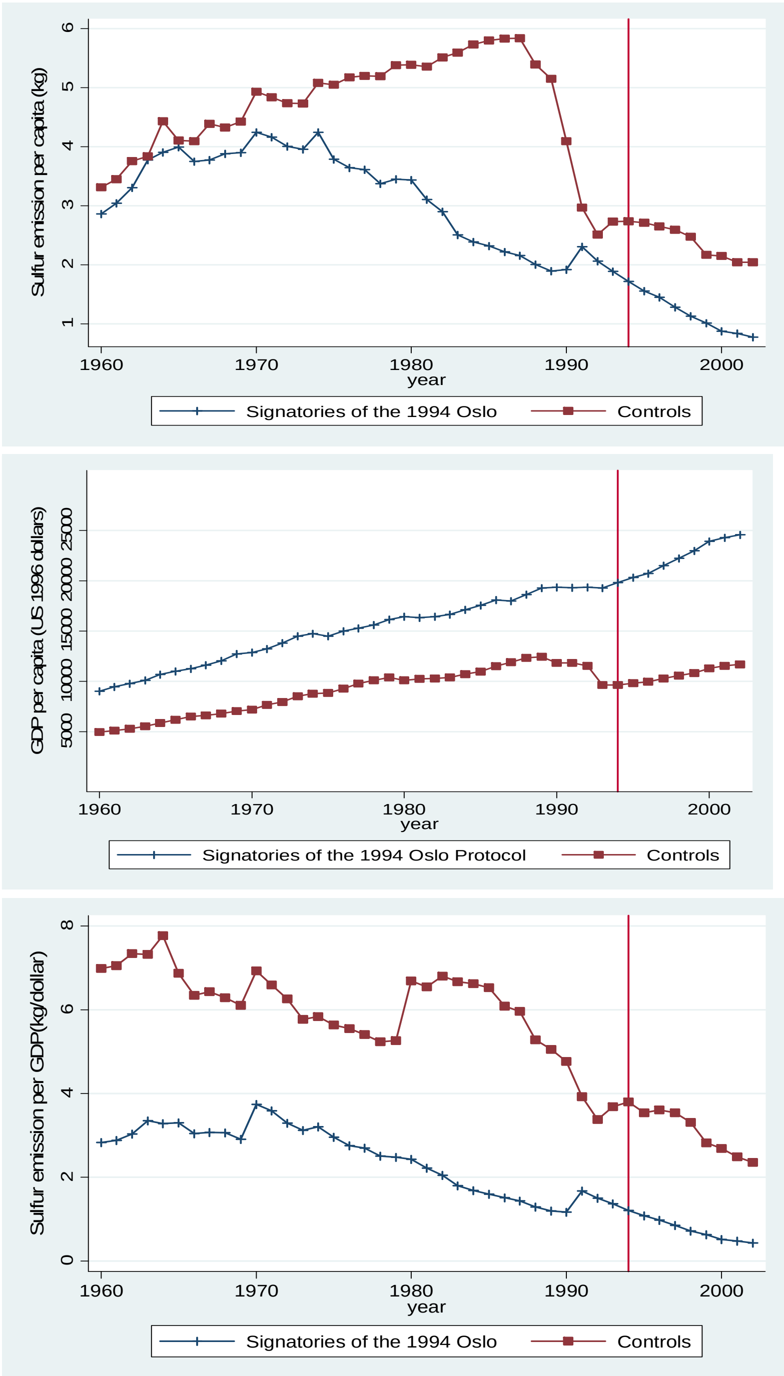
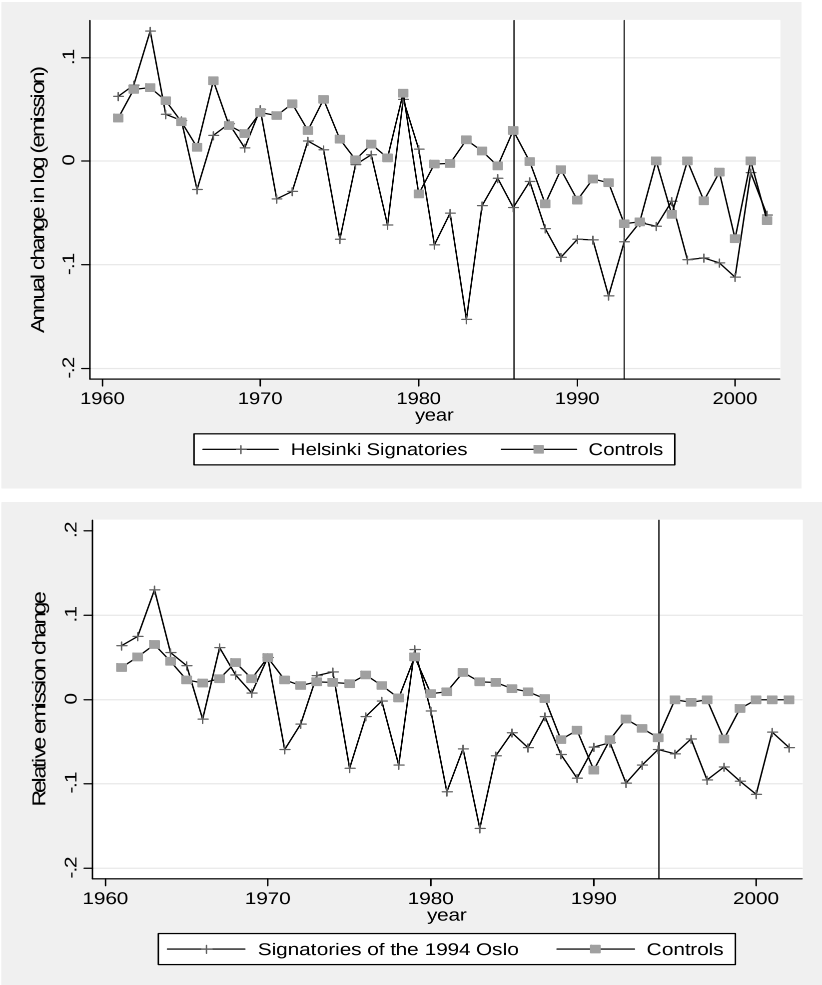
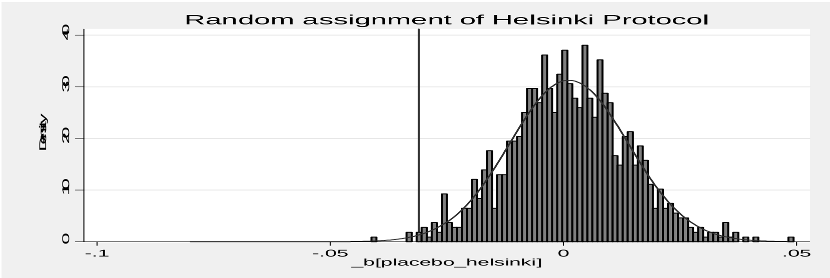


Figure 3. Median annual change in log (emissions) for signatories of the 1985 Helsinki Protocol and the 1994 Oslo Protocols and their respective control groups.

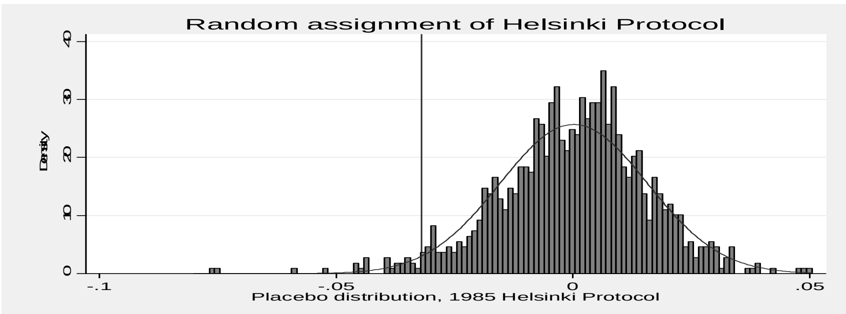


<u>Figure 4.</u> Distribution of effects of placebo protocols for the 1985 Helsinki Protocol and the estimated effect (vertical line) from Model III in Table 2 for different assumptions about the structural draw and the draw.

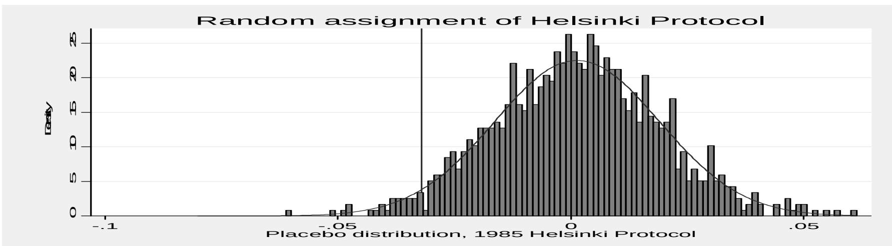
(i) Structural annual draw



(ii) Structural period draw



(iii) Effect parameter restricted to actual period, annual draw



(iv) Effect parameter restricted to actual period, period draw

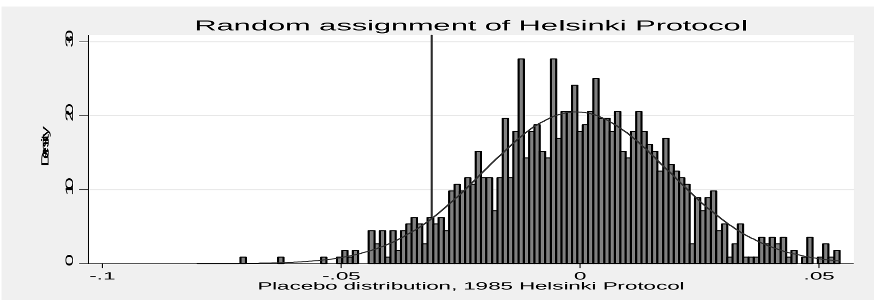
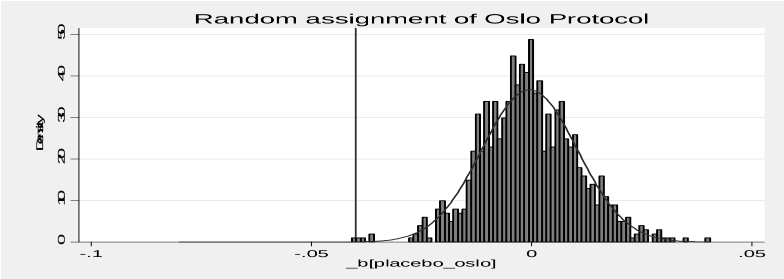
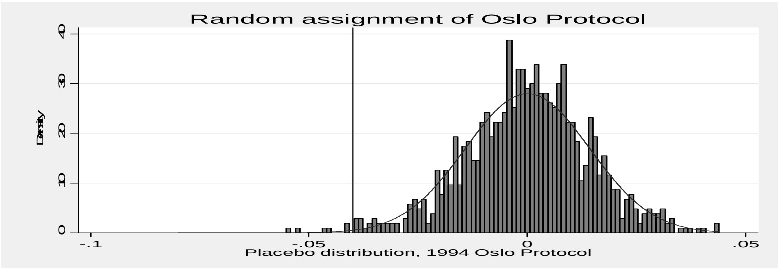


Figure 5. Distribution of effects of placebo protocols for the 1994 Oslo Protocol and the estimated effect (vertical line) from Model III in Table 2 for different assumptions about the structural draw and the draw.

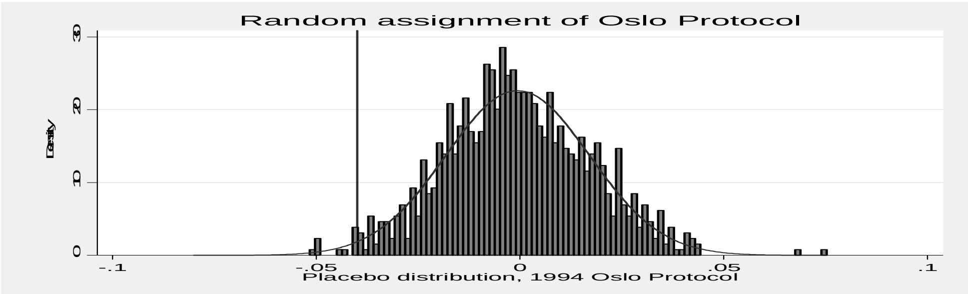
(i) Structural annual draw



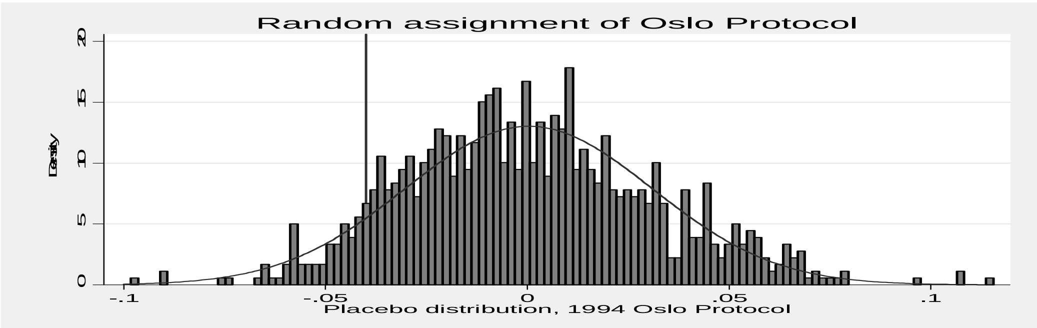
(ii) Structural period draw



(iii) Effect parameter restricted to actual period, annual draw



(iv) Effect parameter restricted to actual period, period draw



Appendix A. Data description

Countries. All countries signed and ratified *The 1979 Geneva Convention on Long-range Trans-boundary Air Pollution.* Since the 1979 Geneva Convention, the political map of Europe has changed. Before the reunification of Germany in 1990, we use West Germany 1960–1989 and East Germany 1970–1989. For the reunified Germany, we use data from 1990–2002. On 1 January 1993, Czechoslovakia underwent a “velvet divorce” into its two national components, the Czech Republic and Slovakia. The two new states inherited the former state responsibilities related to the 1978 Geneva Convention and its follow-up protocols. We use 1960–1990 data for Czechoslovakia and for the Czech and Slovak republics for 1991–2002. In December 1991, the USSR splintered into Russia and 14 other independent republics. We use data for the USSR for 1960–1990 and for Russia for 1990–2002.

Signatories of the protocols. We define a country as a **signatory** of *The 1985 Helsinki Protocol on the Reduction of Sulphur Emissions or their Transboundary Fluxes by at least 30 per cent* if the country signed the protocol and ratified it in 1993. Both Czech and Slovak Republics ratified the protocol in 1993, but Czechoslovakia did not sign. Here we classify Czechoslovakia and its two separated countries as non-signatories of the 1985 Helsinki Protocol. We define a country as a signatory of *The 1994 Oslo Protocol on Further Reduction of Sulphur Emissions* if it signed the protocol and ratified it in 2002.

Sulphur emissions data are from Stern (2006). For 1980–2001, the data are available from the EMEP web site (www.emep.in). Additional data for 1970–1980 are from earlier OECD statistics from Denmark, Finland, France, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom. Stern interpolates data where necessary using Streets and ASL data to derive consistent data. Furthermore, Stern estimates data from other sources for the period 1960–1970 and for non-OECD countries in the period 1990–1980.

For the year 2002, we supplement Stern's data from EMEP (www.emep.in). For Russia 1990–2002, we use EMEP expert emissions predictions and we approximate for the missing years of 1991–1994.

Population data are from PENN World Table Version 6.2 (Heston *et al.* (2006), variable population). For East and West Germany, the USSR, and Czechoslovakia during this time, we use PENN World Table Version 5.6.

Gross domestic product is also from PENN World Table Version 6.2 (Heston *et al.* (2006), We use real GDP per capita in 1996 dollars (variable RGDPCH[16].) For Bulgaria 1980–1990, Czechoslovakia 1960–1990, East Germany 1970–1989, West Germany 1950–1992, and the USSR 1960–1990, we use GDP per capita from PENN World Table Version 5.6. To get consistent data between the two GDP per capita numbers, we regress the GDP per capita from PENN World Table 6.2 on GDP per capita PENN World table version 5.6 controlling for countries and years, and use the predicted GDP per capita in 1996 dollars.

Appendix B. Residual analysis for differential model

Table B.1. Regression model for $\log_e$ emissions, differential model.

Dependent variable: Annual change in $\log_e$ emissions			
	Coefficient	Robust standard error	Standard error
Helsinki 1986–1993	–0.031	0.021	0.020
Oslo 1995–2002	–0.040	0.025	0.025
Gross Domestic Product	0.446**	0.188	0.131
Population	–0.530	1.315	1.251
Transition = 1 after 1990	–0.037	0.023	0.029
Year effects	F = 3.43***		
Country growth effects (g_i)	F = 3.04***		
Constant	0.012	0.027	0.034
R-square	0.269		
N	1,038		

Note: Explanatory variables are annual change in logarithms. Significant levels two-tailed test: * 10 per cent, ** 5 per cent, *** 1 per cent based on robust standard errors.

Table B.2. Analyses of serial correlations for differential models.

Residual (t)	I	II	III	IV	V
Residual (t–1)	–0.032	–0.035	–0.044	–0.039	–0.034
	(0.031)	(0.032)	(0.032)	(0.033)	(0.033)
Residual (t–2)		0.059*	0.056*	0.059*	0.044
		(0.032)	(0.032)	(0.033)	(0.033)
Residual (t–3)			–0.063*	–0.065**	–0.064*
			(0.032)	(0.033)	(0.034)
Residual (t–4)				–0.091***	–0.093***
				(0.033)	(0.033)
Residual (t–5)					–0.041
					(0.041)
N	1,008	978	948	918	888

Note: Significant levels two-tailed test: * 10 per cent, ** 5 per cent, *** 1 per cent.

Table B.3. Correlation matrix for residuals.

	Res(t)	Res(t–1)	Res(t–2)	Res(t–3)	Res(t–4)
Res(t)	1.000				
Res(t–1)	–0.031	1.000			
Res(t–2)	0.046	–0.046	1.000		
Res(t–3)	–0.067	0.058	–0.061	1.000	
Res(t–4)	–0.084	–0.066	0.059	–0.046	1.000
Res(t–5)	–0.040	–0.076	–0.065	0.051	–0.046

Appendix C. Residual analysis for level model

Table C.1. Regression model for log emissions, level model.

Log emission	Coefficient	Robust Stand. Error
(Helsinki 1986–1993)*years since 1986	–0.030***	0.006
(Oslo 1995–2002)*years since 1995	–0.084***	0.013
Log GDP	0.495***	0.010
Log Population	–1.172**	0.584
(Transition)*Years since Transition	–0.049***	0.010
Years since 1960	–0.056***	0.005
Year dummies (1960–2002)	Included***	
Country dummies (30 countries)	Included***	
Interaction term year and country	Included***	
N	1,068	
R-squared	0.988	

Note: Significant levels two-tailed test: * 10 per cent, ** 5 per cent, *** 1 per cent.

Table C.2. Analyses of serial correlation for level model.

Residual (t)	Model I	Model II	Model III	Model IV	Model V
Residual (t–1)	0.791***	0.854***	0.840***	0.826***	0.832***
	(0.018)	(0.032)	(0.032)	(0.032)	(0.033)
Residual (t–2)		–0.075**	0.076*	0.091**	0.077*
		(0.031)	(0.041)	(0.042)	(0.042)
Residual (t–3)			–0.174***	–0.119***	–0.116***
			(0.032)	(0.042)	(0.042)
Residual (t–4)				–0.058*	–0.052
				(0.032)	(0.043)
Residual (t–5)					–0.017
					(0.033)
N	1,038	1,008	978	948	918

Note: Significant levels two-tailed test: * 10 per cent, ** 5 per cent, *** 1 per cent.

Table C.3. Correlation matrix for residuals.

	Res(t)	Res(t–1)	Res(t–2)	Res(t–3)	Res(t–4)
Res(t)	1.00				
Res(t–1)	0.79	1.00			
Res(t–2)	0.59	0.78	1.00		
Res(t–3)	0.38	0.59	0.77	1.00	
Res(t–4)	0.20	0.38	0.59	0.78	1.00
Res(t–5)	0.08	0.21	0.38	0.60	0.79

Department of Economics
University of Bergen
Fosswinckels gate 6
N-5007 Bergen, Norway
Phone: +47 55 58 92 00
Telefax: +47 55 58 92 10
http://www.svf.uib.no/econ