Article 25fa pilot End User Agreement

This publication is distributed under the terms of Article 25fa of the Dutch Copyright Act (Auteurswet) with explicit consent by the author. Dutch law entitles the maker of a short scientific work funded either wholly or partially by Dutch public funds to make that work publicly available for no consideration following a reasonable period of time after the work was first published, provided that clear reference is made to the source of the first publication of the work.

This publication is distributed under The Association of Universities in the Netherlands (VSNU)'Article 25fa implementation' pilot project. In this pilot research outputs of researchers employed by Dutch Universities that comply with the legal requirements of Article 25fa of the Dutch Copyright Act are distributed online and free of cost or other barriers in institutional repositories. Research outputs are distributed six months after their first online publication in the original published version and with proper attribution to the source of the original publication.

You are permitted to download and use the publication for personal purposes. Please note that you are not allowed to share this article on other platforms, but can link to it. All rights remain with the author(s) and/or copyrights owner(s) of this work. Any use of the publication or parts of it other than authorised under this licence or copyright law is prohibited. Neither Radboud University nor the authors of this publication are liable for any damage resulting from your (re)use of this publication.

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please contact the Library through email: copyright@ubn.ru.nl, or send a letter to:

University Library
Radboud University
Copyright Information Point
PO Box 9100
6500 HA Nijmegen

You will be contacted as soon as possible.

4 Research design, variables and data

STEPHAN HEICHEL, KATHARINA HOLZINGER,
THOMAS SOMMERER, DUNCAN LIEFFERINK,
JESSICA PAPE AND SIETSKE VEENMAN

4.1 INTRODUCTION

This chapter gives an overview of the research design of the project, of the variables under investigation and of the data. Given the central research question underlying this project, convergence of national environmental policies is conceived as the dependent variable. Convergence is observed for twenty-four countries over a period of thirty years. Policy convergence is measured as increasing policy similarity over time. Policy similarity is investigated with respect to three dimensions of policy: presence-of-policy, policy instruments and policy settings. In section 4.2, we specify the operationalisation of and data collection for the dependent variable.

As outlined in chapter 3, the explanatory focus of the project is on three factors as the main independent variables that are expected to account for differences in environmental policy convergence: (1) the degree of interlinkage of countries in international institutions with obligatory potential, (2) the degree of interlinkage of countries in international institutions with communicative potential, and (3) the degree of economic interlinkage, i.e., the extent to which a country is connected with other countries by its trade relations. Moreover, we include further explanatory variables (referred to as ‘other variables’ in chapter 3). This is due to the fact that there are several other factors that we conceive as potentially having some impact on national environment policies and also on policy convergence, including cultural similarity between countries, pre-existing policy similarity, political demand exerted by green parties or green movements, the

Figure 4.1 Independent and dependent variables

level of income in a country, as well as environmental problem pressure as a general factor forcing national governments to react through policy measures. Problem pressure indicators used here include carbon dioxide emissions and population density. The central independent and the other variables are introduced in section 4.3 and summarised in figure 4.1.

4.2 DEPENDENT VARIABLE: OPERATIONALISATION AND DATA COLLECTION

4.2.1 *The use of output data*

We measure policy convergence by focusing on output rather than outcome data. We are not interested in the convergence of environmental quality but in the convergence of the regulatory output.

The reason for this choice is that the theories predicting policy convergence have a micro foundation. The theory of regulatory competition assumes that as a result of market forces and industrial lobbying or exit threats, national governments come under pressure to adapt environmental standards and policies. That is, governments are the central actors in this theory. The same is true for policy learning via transnational communication: it is assumed that governments learn from others in international organisations and adapt their policies accordingly. Finally, international harmonisation requires national governments to adjust their policies to the norms agreed upon at the international level. Thus, what counts is what governments (or parliaments, depending on the political system) decide. The theories focus on policy output, not on policy outcomes. Policy output and policy outcome are only very indirectly related. There are many intervening variables affecting the outcomes which are independent of governmental or parliamentary decisions. For example, activities that damage the environment may increase and thus environmental quality degrades more than expected.

Therefore, we use the legal environmental policy output data for the study of convergence. This data type better reflects what is commonly understood as 'environmental policy', namely the 'vast corpus of legislation to restrain citizens and corporations from recklessly fouling the air, land, and water on which they depend for life and health' (Heidenheimer, Heclo and Adams 1990: 309). In that regard, we follow the advice that 'good theories deserve good data' (Baumgartner and Leech 1996) that truly fit the models even if this requires considerable collection efforts.

Policy output data are not systematically compiled by international organisations and thus have to be collected by researchers. Data on policy outcome, instead, can be obtained from existing sources, for example the OECD. Consequently, only a small number of studies employed environmental policy output data in large-n comparative analyses. However, most of these studies remain descriptive (e.g., Binder 2002; Kern, Jörgens and Jänicke 2000; Tews,

Busch and Jörgens 2003), while the rare causal inference analyses have only relied on individual policies (Fredriksson and Millimet 2004; Fredriksson *et al.* 2005; Hironaka 2002). Causal inference studies overwhelmingly employ data on policy outcome, for example emissions or quality data, as proxies for environmental policy.¹

The data collected for this research project imply at least two important innovations. First, the data are not constrained to single measures but attempt to cover a broad range stretching over all environmental media. Second, environmental limit values are used as metrical data expressing the level and strictness of environmental standards.

4.2.2 *Selection of policy items*

Policy convergence is assessed by considering forty different environmental policies that were selected according to several criteria. First, they were taken to form a sample covering all environmental media, ranging from air quality protection to general guiding principles, such as sustainability.

Second, the policies cover the whole range of different policy types that are of particular theoretical interest. Hence, to assess the effects of economic interlinkage (regulatory competition), both trade-related policies (product and process standards) and policies that are not subject to competitive pressures are included. To investigate the effects of the obligatory (international harmonisation) and communicative (transnational communication) potential of international institutions, the policy compilation covers not only so-called 'obligatory' policy items for which a legally binding standard at the international level was introduced at some point between 1970 and 2000, but also 'non-obligatory' measures for which no internationally harmonised rules were in force during the observation period.

¹ Cf. Bernauer and Koubi 2004; Binder and Neumayer 2005; Crepaz 1995; Fisher and Freudenburg 2004; Jahn 1998, 2000; King and Borchardt 1994; Midlarsky 1998; Murdoch and Sandler 1997; Murdoch, Sandler and Sargent 1997; Neumayer 2001, 2003; Ricken 1995, 1997; Scruggs 1999, 2001, 2003; Shandra *et al.* 2004.

Third, the selection of policies was accomplished with the intention of analysing convergence effects with regard to the three different policy dimensions (presence-of-policy, instruments and settings). This way, we are able to analyse whether convergence varies across the different dimensions. Moreover, the inclusion of setting policies allows us to assess the direction of convergence.

The forty policy items selected are listed in table 4.1. The table indicates their characteristics in terms of the three policy dimensions, trade-relatedness, and the presence of international obligations. Of the forty policies, nineteen belong to the group of obligatory international policies in the year 2000, while for twenty-one policies obligatory international measures did not exist throughout the period. In earlier periods fewer policies were obligatory (see table 4.1). Eleven policies are product-related, fifteen are related to production processes and fourteen are not related to products and production processes. Finally, while all forty policies do indicate presence-of-policy, twenty-eight belong to the instrument type and twenty-one to the setting category (see chapter 6.2.2 for details).

4.2.3 Selection of countries and time period

The goal of this research project is to study convergence of environmental policies in Europe. The sample is therefore primarily composed of European countries. From the pool of all European countries, twenty-one were selected. The European group is complemented by three non-European countries. The overall sample thus includes twenty-four countries.

The core of the sample is constituted by member states of the EU. Of the member states in 2000 fourteen are included.² This may look like a lack of variance in the sample with respect to membership. However, the country sample represents all steps in the process of EU enlargements, and therefore over time it includes a sufficient

² The exception is Luxemburg for which data for some, especially the economic, variables are not available separately.

Table 4.1 *List of environmental policies for the dependent variable*

Policy	Environ- mental media	Policy type			Trade relatedness of policy				Obligatory policy			
		Setting	Instrument	Policy presence	Product	Process	Non- trade related		1970	1980	1990	2000
1 Sulphur content in gas oil	Air	•	•	•	•				•	•		•
2 Lead in petrol		•	•	•	•				•	•		•
3 Passenger cars NOx emissions		•	•	•	•				•	•		•
4 Passenger cars CO emissions		•	•	•	•			•	•	•		•
5 Passenger cars HC emissions		•	•	•	•			•	•	•		•
6 Large combustion plants SO ₂ emissions		•	•	•		•				•		•
7 Large combustion plants NOx emissions		•	•	•		•				•		•
8 Large combustion plants dust emissions		•	•	•		•				•		•

Table 4.1 (*cont.*)

Policy	Environ- mental media	Policy type		Trade relatedness of policy				Obligatory policy			
		Setting	Instrument	Policy presence	Product	Process	Non- trade related	1970	1980	1990	2000
9	Coliforms in bathing water	•	•	•			•		•	•	•
10	Hazardous substances in detergents		•	•	•				•		•
11	Efficient use of water in industry			•		•					
12	Industrial discharges in surface water lead	•	•	•		•					
13	Industrial discharges in surface water zinc	•	•	•		•					
14	Industrial discharges in surface water copper	•	•	•		•					

[illegible]

Table 4.1 (*cont.*)

Policy	Environ- mental media	Policy type			Trade relatedness of policy				Obligatory policy			
		Setting	Instrument	Policy presence	Product	Process	Non- trade related		1970	1980	1990	2000
25 Noise emissions standard from lorries	Noise	•	•	•	•				•	•	•	•
26 Motorway noise emissions		•	•	•			•					
27 Noise level working environment		•	•	•		•			•	•	•	•
28 Electricity from renewable sources	Resources		•	•		•						
29 Recycling construction waste				•		•						
30 Energy efficiency of refrigerators	Climate		•	•	•							•

[illegible]

number of non-EU countries. The EU founding members – Belgium, France, Germany, Italy and the Netherlands – form the core group of the sample. Hence in 1970, only five countries out of twenty-four are subject to EU legislation. By 1980, Denmark, Ireland and the UK had become EU members, and by 1990, the Southern enlargement countries, Greece, Portugal and Spain, had joined the EU. At that time, EU members constitute still less than half of the sample (eleven). Former members of the European Free Trade Association (EFTA), Austria, Finland and Sweden, joined the EU in 1995. The group of EU members in 2000 is complemented by two groups of countries from CEE: with Hungary, Poland and Slovakia three countries are included that became EU members in 2004. Finally, Bulgaria and Romania belong to the sample that will probably accede in 2007. Two further countries complete the European group, the EFTA member states Norway and Switzerland.³

To control for the EU harmonisation effects and to increase variance regarding several crucial covariates, the sample includes three non-European countries. The US and Japan are countries with a high level of economic development and a high trade volume, while Mexico shows a comparatively low level of economic interlinkage and economic development.

The period of observation reaches from 1970 to the year 2000. This longitudinal perspective ensures that the data on environmental policy are more than a current snap-shot, and that policy changes can be studied in the long term; this is indispensable to each study that aims to analyse processes of convergence. The year 1970 is chosen as a starting point because at that time environmental policies began to develop at the national, as well as at the international, level

³ Several other European countries had to be omitted from our sample. The Czech Republic is not included, as only one successor state of former Czechoslovakia should be part of the sample. This is to avoid an artificial bias towards convergence, because these countries had the same policies until 1990. The same applies to Germany, where only the Western part is included for the period before 1990. For similar reasons, as well as for problems of data accessibility, no country from the former Soviet Union (the Baltics) or former Yugoslavia is part of the sample.

(e.g., with the 1972 UN Conference on the Human Environment in Stockholm). The observation period ends in 2000, the last year for which data were available for all variables when data collection started in 2003.

Basically, there are two extreme strategies for analysing convergence: first, one can take only two points of measurement, and second, one can use a time series design. Taking only two points in time (1970, 2000) does not allow us to see intermediate developments. Many policy changes are likely to have happened in between. Time series design is not adequate either, because of low variance over time: environmental policy is not expected to change on a yearly base – the process of domestic policy-making takes its time, as does the influence that comes from the international level. As it is rarely the case that several changes occur within a decade, an intermediate strategy is employed: policy convergence is measured for each decade separately. The research design is therefore based on the measurement of policy similarity at four points in time: 1970, 1980, 1990 and 2000. This way most policy changes are covered.

4.2.4 Data collection and processing

The dependent variable describes the evolution of forty environmental policy items in twenty-four countries over thirty years. This vast amount of data was collected with the help of expert surveys. This way, a comprehensive database was compiled, offering unique opportunities for studying patterns of environmental policy convergence between the countries involved. In the following, we discuss the measures taken to ensure data validity and reliability during data collection and data processing.

Expert surveys The necessary output data could not be compiled from existing sources. The ECOLEX database on environmental law, developed by the World Conservation Union, the UN Food and Agriculture Organisation and the UN Environment Programme comes closest to our requirements, but even here ‘historical’

information on early environmental legislation is often lacking. To obtain this information, it proved necessary to access the original laws, decrees and ordinances.

In so doing, we relied on expert surveys. Expert surveys are generally seen to have several crucial advantages over alternative approaches. Precisely because they reflect the judgements of experts, they acquire a certain weight and legitimacy. Furthermore, as expert judgements are reasonably comprehensive, they permit the collection of comparable and standardised data. Despite these advantages, however, problems of validity and reliability of the data may occur. For instance, the same question may be interpreted in different ways by different experts. Measurement errors could occur for example when experts overlook legislation. However, as we asked the experts for facts (introduction of laws, use of instruments, setting of limit values etc.) and not for opinions or attitudes, most of the problems regarding the validity and the reliability of data that are common to social science surveys did not apply here.

An important step in guaranteeing the quality of data was to make a careful selection of experts to carry out the work. First, in order to address adequately the wide variety of questions in the survey, the experts had to have a comprehensive overview of environmental policy development from 1970 to 2000. Second, the group of experts should be as homogeneous as possible, so as to ensure the comparability of data (Frankfort-Nachmias and Nachmias 1996). In order to meet both criteria, we limited our selection to academics with a background in political science or law and preferably working at research institutes or universities.

The design of the questionnaire and the manual The questionnaire had a standardised set up so as to limit the room for interpretation to a minimum (see table A4.1 in the annex for an example). At the same time, open questions ('remarks section') were included to allow for the description of item-specific characteristics. This facilitated the checking of data comparability. The questionnaire was accompanied

by a detailed manual, presenting general instructions and definitions as well as background information on each individual policy item to ensure the validity and reliability of the data (see below).

As set out above, the selection of policies addressed by the project covered one or more of the three dimensions: presence-of-policy, policy instrument and setting. The first type – ‘presence-of-policy’ questions – asked for the mere existence of certain environmental policies in 1970, 1980, 1990 and 2000 and, if they existed, the year of their introduction.

The second type – ‘policy instrument’ questions – addressed the kinds of instruments used to tackle environmental problems as well as their application over time. Experts were asked to select one ‘dominant’ instrument type from the following list:

- (1) obligatory standard, prohibition or ban
- (2) technological prescription
- (3) tax or levy
- (4) subsidy or tax reduction
- (5) liability scheme(s)
- (6) planning instrument
- (7) public investment
- (8) data collection / monitoring programmes
- (9) information based instrument
- (10) voluntary instrument
- (11) other instruments⁴

Since in practice almost always mixes of instruments are used to address a given problem, this implied that experts assessed the role of different instrument types before selecting the ‘dominant’ one. In order to prevent an arbitrary interpretation of what is ‘dominant’ and due to problems the pre-testers encountered with this question,

⁴ The last instrument type served only as a residual category in case an expert felt unable to assign the national instrument to one of the existing types. For the rare cases where this occurred we were usually able to subsume the national policy instrument under one of the existing categories after consulting the expert.

it was specified in the accompanying manual that if there was an instrument mix including an 'obligatory standard, prohibition or ban', this instrument should be counted as dominant. This instruction rests on the argument that obligatory instruments tend to dominate over other instruments in the mix, because other instruments like subsidies or information rights are then usually designed to support this obligation. Although this induced a certain bias towards the 'obligatory' type of instruments, this procedure ensured that experts interpreted the question in a consistent way and that results would be comparable. In addition, experts were requested to indicate the exact name of the instrument and briefly explain its function (e.g., gasoline tax reduction for lead-free petrol as an instrument to regulate lead emissions from vehicles). This was to be accompanied by the name of the legislation containing the regulation in order to facilitate cross-checking by the project team.

The third type of questions related to environmental policy settings. Here experts were asked to provide numerical values of limit values, emission standards, quantified goals or tax rates. In addition, the experts were requested to indicate the unit of measurement. Finally, each policy item in the questionnaire contained a section for further remarks and one for references. The remarks section enabled experts to describe in more detail particular characteristics of the policy in question. This enhanced validity as it enabled the research team to retrace the decisions taken by the expert in problematic cases. The reference section improved reliability by enabling the project team to verify the information.

The aim to receive adequate, complete, precise and comparable data was further served by providing an extensive manual to accompany the questionnaire. In the first part of the manual definitions of basic terms and instructions were given that allowed for a consistent interpretation of the questions. The second part contained sections on every policy item covered by the questionnaire, clarifying technical terms used and defining very precisely which information we were looking for (see table A4.2 in the annex for an example).

Before the questionnaires were sent out, two experts were asked to carry out a pre-test. Pre-testing is the only way to evaluate in advance whether a questionnaire causes problems for interviewers or respondents (Graesser *et al.* 2000: 459; Presser 2004; 109). The comments of the two pre-testers led to substantial improvements of both the questionnaire and the manual.

Experts were asked to keep in touch with the research team during the entire process of filling in the questionnaire. If further specifications of questions turned out to be necessary, these were immediately communicated to all other experts. This further reduced the risk of inconsistencies and increased the validity of the data.

Data processing After the questionnaire had been received back from the experts, two further steps were undertaken to increase the validity and reliability of the data. First, the material provided by the experts was systematically cross-checked. For this purpose, we relied on data available from other databases, for instance, the OECD, the European Environment Agency, as well as from ECOLEX and other relevant scientific literature, including an existing data set compiled by one of the project partners (Busch and Jörgens 2005a). Moreover, a random sample of six questions for detailed cross-checking was selected for each country, which took into account the variety of questions asked (presence-of-policy, instruments, settings). All checks were carried out in order to see whether the information given by the experts was reliable.

Second, the data set was 'cleaned' by clarifying and correcting remaining inconsistencies and errors. Three basic types of inconsistencies emerged:

- (1) Different units of measurement. In the large majority of cases this only entailed the recalculation of settings into a single unit, e.g. mg/m^3 . In a few cases, however, differences were not due only to the application of different measurement units as such, but also to the use of different measurement procedures. Emissions,

for instance, were measured at different heights or at different distances from the source. If possible, transformation tables were created. Where this was not possible, which was the exception rather than the rule, this constituted a missing value in the data set. In those cases, of course, the existence of a limit value was still recorded to avoid unnecessary loss of data.⁵

- (2) Differentiated standards. For a limited number of items, it turned out that standards in some or all countries were even more differentiated than anticipated in the questionnaire and manual, for example different limit values for different industrial (sub-) sectors, for different sizes of plants, for old and new plants, for different regions. In those cases, as a rule, we chose the environmentally strictest values in place. For example, for noise levels around streets, we took the values applying at night in the highest protection zones.
- (3) Different interpretations of policy instruments in use. In a few cases, the precise character of policy instruments in use turned out to be difficult to determine for experts, leading to inconsistent interpretations. In relation to taxes on water extraction, for instance, it proved impossible in most countries to distinguish between a tax meant for regulatory purposes (i.e., as an incentive to change behaviour) and a tax only intended for covering administrative costs, or a combination of both. Therefore, it was decided that all taxes/fees/charges given in questionnaires should count as a tax.

The phase of data processing was completed with the compilation of a raw data set in SPSS statistical software. Presence of a policy for each point of measurement was coded as '1' (yes) or '0' (no). Policy instrument data were re-coded on a scale ranging from '1' ('obligatory standard, prohibition or ban') to '10' ('voluntary instrument'). Data on settings were included as numerical values. The final result was

⁵ In the case of air pollution from passenger cars, emission standards were based on different test cycles and harmonisation was not possible in the case of CO emissions. In order not to lose data and guarantee comparability, two sub-groups were constructed, with limits expressed as g/km and volume per cent, respectively.

one complete set containing the entire information for the sample countries from 1970 to 2000, as well as four separate data blocs for 1970, 1980, 1990 and 2000.

4.2.5 Types of convergence

It is essential to any study of convergence to address the question of how convergence can actually be measured. The theoretical contributions by Bennett (1991: 296ff.), Seeliger (1996: 288) and Unger and Waarden (1995: 3) have convincingly demonstrated that convergence means, by definition, the increase of policy similarity. From the literature, different operationalisations of this basic definition of convergence can be deduced which refer to different ideas of similarity (see Heichel, Pape and Sommerer 2005).

Sigma-convergence as ‘growing together’ The most basic way of assessing policy convergence is to analyse the extent to which the policies of countries have become more similar over time (sigma-convergence). Following this approach, convergence occurs if there is a decrease in variation of policies among the countries under consideration. This concept is named after the algebraic notation for variance, sigma (Sala-i-Martin 1996: 1326ff.). In its traditional form, a decreasing coefficient of variance, the normalised standard deviation, also known as dispersion index, indicates convergence. However, many other ways to operationalise sigma-convergence are applied (see Heichel, Pape and Sommerer 2005). This approach to convergence is used in studies on tax revenue and governmental expenditure or on tax rates as well as on trade policy. As the most important approach to convergence, sigma-convergence will be analysed in depth in chapters 5 and 6.

Beta-convergence as ‘catching up’ A second approach to investigating similarities across countries, often used in economics, is the study of beta-convergence. It occurs when poor economies grow faster than rich ones and it is named after the growth coefficient, beta

(Sala-i-Martin 1996: 1326ff.) An example of an adequate transformation from economics to the context of analysing policy lags and policy frontrunners in immigration policy shows the valuable contribution of this logic to policy analysis (Holzer and Schneider 2002: 64). Here, beta-convergence occurs when the degree and speed of tightening asylum regulations in formerly liberal countries is higher than in formerly restrictive ones. For environmental policy, beta-convergence occurs when countries with lax environmental regulations tighten their regulations faster than those applying strict regulation from the beginning. This concept will be applied to the assessment of convergence of environmental policies in chapter 5.

Gamma-convergence as ‘mobility’ A recent variant of convergence research is gamma-convergence. It was developed in response to beta-convergence, which has been criticised for not sufficiently capturing aspects of cross-country dynamics (Boyle and McCarthy 1999). For the analysis of gamma-convergence, country rankings for different points in time are compared to assess the mobility of countries. If countries in the first ranks fall behind or countries in the last ranks catch up over time, convergence occurs. Policy change is analysed by simple measures of association, like the Kendall index of rank concordance (Boyle and McCarthy 1999). A low degree of similarity between rankings indicates high mobility of countries over time. The concept of gamma-convergence adds an additional perspective to the study of policy convergence as it may occur where other approaches do not detect changes. Country rankings, for example, may change without a significant decrease of cross-country variation or without an approximation towards an exemplary policy model. Although the study of gamma-convergence is not very common yet, it represents nonetheless a promising tool for policy studies, as the idea of comparing ordinal classifications is also compatible with a qualitative research design. In addition to changes in growth rates, gamma-convergence will be studied in chapter 5.

Delta-convergence as ‘minimising distance to an exemplary model’

Many studies on convergence analyse different aspects than those sketched above. What they have in common is an approach to convergence based on the decreasing distance of policies towards an exemplary model, for example a model promoted by an international organisation or a frontrunner country. Based on the algebraic notation of ‘distance’, delta, this concept can be called delta-convergence. Empirically, gamma- and delta-convergence often occur simultaneously. If countries reach total similarity relative to a policy model, variance between them is obviously reduced. However, if delta-convergence is not complete, policies may approach the model by parallel moves without becoming more similar. By the reference to an exemplary policy model that may be the strictest, the laxest, or the mean level of environmental regulation, delta-convergence adds aspects of the direction of convergence to the analysis. It will be applied in depth to environmental policy convergence in chapter 7.

Depending on the type of convergence investigated, empirical results might be interpreted very differently. Evidence of sigma-convergence, for instance, does not necessarily mean that there is also gamma-convergence. Moreover, evidence of beta-convergence does not imply that there must also be sigma-convergence: the fact that laggard countries change more fundamentally than leader countries is not a sufficient condition for a decrease in variance across all countries.

4.3 INDEPENDENT VARIABLES: OPERATIONALISATION AND DATA COLLECTION

4.3.1 *Institutional interlinkage*

In chapter 3, a central role for the explanation of convergence is assigned to the influence of integration into international institutions. The operationalisation of institutional influence is thus of great importance for the analysis of convergence. There is no standard tool for the operationalisation. The effects of international institutions on domestic policy change are only sporadically included in quantitative

empirical studies, mainly through the inclusion of membership in certain international organisations like the EU or the OECD.

Two different mechanisms are supposed to effect policy convergence through international institutions: international harmonisation and transnational communication. Therefore, two variables will be developed and later tested separately: the obligatory potential and the communicative potential of institutional interlinkage.

Although differentiated according to the respective mechanisms, both variables are based on national membership data. This is necessary, as the data on international institutions has to be linked to the national level at a certain point: policy convergence is measured at the level of countries or country pairs. As a first step of operationalisation of institutional interlinkage, a 'membership' variable is calculated based on the membership of countries in thirty-five selected international institutions. Second, the membership data are weighted according to the institutions' 'encompassingness', as international institutions differ in size and range of activities. Third, two indicators of the obligatory and communicative potential of an international institution are introduced, by which the 'membership and encompassingness' variable is weighted. This provides us with the two indicators for the mechanisms of international harmonisation and transnational communication. Altogether, we thus develop four different variables of institutional interlinkage that will be used for different purposes in the quantitative models.

Membership in international institutions As outlined above, this project attempts to deliver a comprehensive view of the interactions in the field of European environmental policy. This is not just a matter of a wide-ranging selection of policies for the dependent variable; it should also represent the various influences at the level of international cooperation and coordination in the field of environmental policy. Therefore, all relevant institutions are included in an aggregate variable of institutional interlinkage in the environmental field (see annex table A4.3).

Altogether thirty-five international institutions were selected. The selection of institutions is based on the *Yearbook of International Cooperation on Environment and Development* (2002), complemented by some other institutions. Only those institutions which have at least three contracting parties are included, and at least two countries from our sample should be members of these institutions. This excludes the vast amount of bilateral agreements – including them would have been very time consuming without adding much variance.

The institutions listed in table A4.4 in the annex can be classified into five different categories. The first category involves political institutions with a very broad field of activity (general multi-issue organisations). They are of major importance for international relations in general, and, among others, possess a department for environmental policy. Examples are the EU and the OECD. A second group consists of institutions that are predominantly occupied with environmental policy, for example the UN Environmental Programme (UNEP) or the International Union on the Conservation of Nature (IUCN) (environmental multi-issue organisations). Thirdly, there are international institutions that have their main focus on a different field but include environmental aspects as policy tasks, such as the WTO or the Organisation for Security and Cooperation in Europe (OSCE) (non-environmental single-issue organisations). Global or European environmental regimes are a fourth category of institutions (environmental single-issue organisations). An example is the Vienna Convention for the Protection of the Ozone Layer of 1985. Finally, regional regimes like the Conventions on river protection or the Alpine Convention were included.

Encompassingness The aggregation of institutional scores necessitates some sort of weighting. Differences in size and relative importance in international relations are considerable, for example, between the EU and single-issue regimes like the Espoo Convention. To differentiate for size effects, we develop an index for the

encompassingness of international institution that is used as a weighting factor.

The weighting factor for encompassingness includes three indicators that are supposed to represent various dimensions of size effects. First and foremost, the importance of an international institution is measured by the breadth of its fields of activity. On the one hand, this represents the relative size of an organisation in general, not only for environmental policy; on the other hand, it is important to note that environmental protection is a cross-sectional policy, hence other policy fields are of importance, as well. We count the number of policy areas, assuming that the level of encompassingness increases with the number of policy areas an institution includes. An almost nation state-like jurisdiction, such as the EU, constitutes the top of the scale. The scheme for the classification of fourteen policy areas is borrowed from the structure of the General Directorates of the European Commission. The original scale from one to fourteen is transformed into a simpler one with three categories only: more than ten areas, five to ten areas, and fewer than five areas (see table A4.3 in the annex).

This first indicator represents the influence of an international organisation or regime only in very general terms, and it is obviously not sufficient. Therefore, a second indicator is included that reflects the scope of environmental problems that are regulated by this institution. They are categorised into three different classes: the institution deals either with all aspects of environmental policies, with various issues, or with one environmental issue only, such as the Convention on Biological Diversity. This indicator increases the difference between encompassing international organisations and single-issue regimes.

The final indicator of encompassingness refers to differences in the relative importance of environmental aspects for an institution. Although, for example, the scope of environmental issues addressed by the WTO is broader than for environmental regimes, it is obvious that environmental protection is not the main subject

in this organisation. Here, we distinguish between three different degrees of priority of environmental problems: in an international institution, they can play a dominant role, be an issue amongst others, or be of no special importance at all. This indicator leads to a downgrading of institutions like the WTO that are large and influential in general, but not in the field of environmental protection.

For the index of institutional encompassingness, these three indicators are summed up and comprised into a weighting factor with a scale from 0–9 that is normalised to 0–1 (see table A4.3 in the annex).⁶ The value for each institution is given in the ENC column of table A4.4 in the annex.

Date of ‘signature’ v. ‘entry into force’ Across all institutions, we collected two different types of membership data, according to differences in the indices for obligatory and communicative potential that will be introduced below as further weighting factors for the data on national membership. The date of accession is operationalised differently for the two factors. For obligatory potential, we apply the date of ‘entry into force’ because a treaty is not binding for a country before the national ratification process has been finished. For communicative potential, the date of ‘signature’ is used, since we assume that transnational communication starts even before an international agreement is ratified. For some institutions, the difference between ‘signature’ and ‘entry into force’ is up to five years. All institutions are captured according to that scheme except for one: for the EU only one date is collected, that is the beginning of accession talks. We assume that the process of taking over the *acquis communautaire* starts already with the beginning of accession talks.

Obligatory potential International harmonisation will be tested by adding an index of obligatory potential (OP in table A4.4 in the annex)

⁶ The reliability for the scale of the index is confirmed by a test of Cronbachs α .

to the data on institutional membership and encompassingness. The obligatory potential of an institution strongly depends on the legal status of its regulatory output. In a first step, this output can be categorised by its degree of bindingness for domestic policy-makers. Categories for bindingness can be deduced from common categories of international law (Abbott *et al.* 2000), leading to a scale with three different types of regulatory output results.

The highest degree of obligation in the sphere of international legislation emanates from supranational law. Even though in international law studies its distinction from other types of law is not unchallenged (cf. Bogdandy 1999; Neyer 1999; Zürn 2000), this category finds wide support. Regulations of this type are implemented prior to national law and can also be enforced by national courts or specific international courts; regulatory output is directly binding without a process of ratification through national parliaments. In supranational law, one finds three different types of output. Determinate and interpretable regulations, represented in EU legislation by regulations and directives, are complemented by the category 'decision', which is also known from EU regulations and comprises regulations with a limited or individual scope of application, such that the binding effect on policy-making in general will be lower. Decisions of the WTO dispute resolution panel also belong to this category.

International hard law is the second most binding category. This is the traditional category of international law on the basis of multilateral contract law. It is also binding, but not as directly as in the case of supranational law. In international hard law, we mainly find international treaties or protocols that have to be ratified by the national parliaments. With respect to their legal bindingness the latter are not equal to treaties. However, protocols like the Kyoto protocol to the International Framework Convention on Climate Change are instruments for interpretation of treaties. Thus, they rank lower in their obligatory potential, but are definitely more obligatory

than international soft law.⁷ Examples for international hard law are found in international conventions, such as the 1979 Convention on Long-Range Transboundary Air Pollution.

The third category comprises international soft law; any regulatory output of an international institution that does not explicitly fall under the first two categories is subsumed under this category. Thus, a wide range of different regulatory output belongs to international soft law.

First, we find resolutions that are decisions of the institution or of one of its organisational bodies with the intention to bind, but not in the form of international contract law. Second, and less binding, there are coordinating measures and third, there are recommendations. Finally, there are measures like benchmarking or opinion statements.⁸ The scale for 'type of law' is an additive one, as different types of regulatory outputs exist in parallel and complement each other. It is calculated as a weighted sum, resulting in a scale from 0 to 6 (1+2+3; see table A4.3 in the annex).

Whereas type of law is an obvious indicator for an institution's obligatory potential, it is not the only one. A second indicator is used to differentiate between institutional capacities to monitor the compliance of its members. Four categories are distinguished, defined by different types of monitoring found in international institutions. Monitoring by courts is the strictest category (e.g., the European Court of Justice). Specialised monitoring bodies, which fall into the second category, are found in many international institutions (e.g., the Implementation Committee for the Vienna Convention). Such a monitoring body is not equivalent to a court; however, it

⁷ According to the Vienna Convention on the Law of Treaties from 1969, the disregard of international hard law has severe consequences. Also, by definition of the UN Charter, hard law gains its obligatory potential from the possibility of settling disputes before international courts; in addition, reprisals are legitimate for other members, if a 'black sheep' fails to comply with the rules.

⁸ The bindingness of international soft law is defined by the fact that against a non-complying member, the instrument of retortion as an unfriendly measure of self-help as a response to mis-conduct is allowed according to the UN Charter.

also controls the efforts to implement international law. The third category comprises reporting duties found in many international regimes. An example is the UN Framework Convention on Climatic Change, where national reports are obligatory for national governments. The weakest monitoring instrument is found in the classical institution of diplomacy that can also exert pressure on domestic policy-makers where no other means are available at the international level.

Both indicators are combined into an index of obligatory potential with a scale from 0 to 10, where the type of law has a slightly higher weight (see table A4.3 in the annex).⁹ To build a variable representing the effects of legal obligation through institutional inter-linkage, this index for all thirty-five institutions is used as a second weighting factor (after encompassingness) for the institutional membership data on country level.

Communicative potential To measure effects of transnational communication, we develop an index of communicative potential (CP in table A4.4 in the annex). It measures the influence of international institutions that goes beyond effects of international law and supranational legislation. Four indicators are selected to cover the relevant dimensions of communication in international institutions.

The first indicator is frequency of interaction. It captures how often members of national governments or ministerial bureaucracies meet in international institutions. The communicative link between the international and the national level is very direct in this case, as the actors themselves are involved in national policy-making. The data on the frequency of meetings have been gathered for all organisational bodies that are concerned with environmental issues, such as committees and working groups on environmental policy, as well as for all organisational bodies that are higher in the hierarchy than these former ones, such as a conference of parties or a general council;

⁹ The reliability for the scale of the index is confirmed by a test of Cronbachs α .

there, environmental aspects may also be discussed among other issues. Unit of measurement is the number of meetings per year. Of all organisational bodies, the one with the maximum frequency of interaction is chosen, as it best represents the potential of informational exchange. The original scale for the number of meetings (0.33–150) is transformed to a scale from 0–10 (see table A4.3 in the annex).

The information on frequency of meetings is complemented by the number of organisational bodies where members of national governments or ministerial bureaucracies meet on environmental issues or on general issues. There may be cases where the maximum frequency of interaction is low, whereas the actual communicative potential is higher due to the presence of many organisational bodies. As for frequency of interaction, only those organisational bodies are included where representatives of national governments meet. The original scale already reaches from 0 to 10; thus, it does not have to be transformed.

The third indicator captures the extent to which national governments are represented permanently in international institutions or not. The communicative potential of an institution is supposed to be higher if there is a permanent delegate coordinating the informational exchange between international institutions and national governments. Presence of delegates is simply measured 'yes/no' and finally transformed to the same normalised scale as the other indicators for communicative potential.

The fourth indicator, size of staff, includes information on the professional staff permanently working for international institutions. The more employees working for an international organisation, the larger the amount of informational output this institution is able to spread, even without the initiative of national governments. The number of permanent employees in environmental departments reaches from 0 to 574, which is again transformed to a scale from 1–10.

These four indicators are aggregated to an index of communicative potential, with a scale from 0–40 (see table A4.3 in the

annex).¹⁰ As for obligatory potential, the institutional scores for communicative potential are used as weighting factors for membership data, in addition to information on encompassingness of institutions. As mentioned above, membership data for the communicative potential variable differ from obligatory potential, as the relevant point of reference for a country's accession is 'signature', not 'entry into force'. Furthermore, information on the length of membership is included in the membership data, as we assume that communication through international institutions works more efficiently in long-established networks than for newcomers.

Data collection The basic data needed with regard to membership of individual countries in individual institutions over time were taken from the *Yearbook of International Cooperation on Environment and Development*, published annually by the Fridtjof Nansen Institute (FNI) in Oslo, which provides a comprehensive collection of international environmental organisations. In addition, websites of institutions and national governments were used.

Further data were needed for weighing institutions according to their encompassingness, their obligatory and their communicative potential. This information was gathered mainly by consulting and comparing websites of the respective institutions, as well as by studying the wording of the relevant treaties. In cases where the information provided by the websites was incomplete or unclear, people directly involved in those institutions were asked.

The institutional data, with the exception of membership over time and type of law, were collected only for 2000, because data for earlier years were not available for all indicators. For some of the indicators, it may be safely argued that substantive change over time is not very likely, for example monitoring capacity. For other indicators, for example the number of contacts with national

¹⁰ The reliability for the scale of the index is confirmed by a test of Cronbachs α .

representatives, development over time is more likely to occur. However, the overriding trend is one of increasing international institutional involvement around environmental issues between 1970 and 2000. Therefore, even if individual institutions have followed this trend somewhat more rapidly than others, it may be assumed that the overall *relative* picture among our thirty-five institutions has remained more or less stable.

4.3.2 *Economic interlinkage*

The classical approach to operationalising economic interlinkage is trade dependence. The intensity of an economy's integration with the rest of the world can be measured by trade flows. The variable conventionally used is trade openness, i.e., trade volume weighted by the size of a country's GDP (e.g., Li and Reuveny 2003). With regard to trade, we use figures on exports and imports of goods only, as imports and exports of services are not related to competitive pressure on products or production as a result of environmental regulation. For each country and each point in time, we use the rather common formula of trade figures as a percentage of total GDP (see table A4.3 in the annex).¹¹

While the index of trade openness will be used for the analysis of delta-convergence in chapter 7, for the analysis of sigma-convergence by the pair approach (chapter 6), a different version of trade openness is applied. The pair approach is based on the comparison of each country pair in the sample. Thus, the data used have to be based on bilateral trade figures. These allow the competitive situation between two national economies to be captured in a more

¹¹ See, for example, Alesina and Wacziarg 1998; Garrett and Mitchell 2001; Hallerberg and Basinger 1998: 334; Korpi and Palme 2003; Li and Reuveny 2003; Rudra 2004; for a concise discussion see Dowrick and Golley 2004. We rely on this approach, although some authors work with trade/PPP GDP (GDP expressed in purchasing power parities). This measure of 'real openness' (Alcalá and Ciccone 2004), however, is criticised with convincing arguments by Rodrik, Subramanian and Trebbi (2002: 18) for being strongly correlated with income levels.

direct way. The trade volume is weighted by the size of the smaller national market of a country pair in order to represent the degree of dependence in bilateral trade relations. To operationalise further conditions that can be derived from theories of regulatory competition (chapter 3), bilateral trade openness is corrected by the existence of integrated market economies: if one of the countries is not a member of GATT/WTO and does not have a market economy, there is no competition between a country pair (for the formula see table A4.3 in the annex). Both types of data could be easily obtained from OECD and World Bank sources.

4.3.3 *Other variables*

Cultural similarity is operationalised as the sum of three different indicators: (1) the existence of a common border between a country pair,¹² (2) the sharing of a common language and (3) the existence of common historical and religious tradition, expressed by the similarity of religion between two countries.¹³ Hence, this indicator can vary between 0 and 3 for each country pair (see table A4.3 in the annex). We conceive this variable as static, which means that cultural ties and geographic proximity are assumed to be highly stable factors that are not subject to far-reaching changes over time.

Economic development is based on the income figures measured as GDP per capita (see table A4.3 in the annex). In the pair approach the income figure is multiplied by the difference in income between the two countries.

Domestic political demand for a comprehensive and stringent environmental policy could possibly be expressed in several ways, such as the number of environmental organisations or the public

¹² Sea borders count, if located within a 200-mile zone. This results, for example, in a common border between the Netherlands and the UK or Denmark and Sweden.

¹³ All languages and all dominant religions count. Because of the pairwise-comparison, countries with more than one language or dominant religion cause no problems. Whenever two countries have a language or dominant religion in common they are rated similar.

opinion data on environmental awareness. In view of their important role in bringing about political changes in domestic environmental policy, we concentrate on the political influence of national green parties. Green parties' political influence can also be seen as a very general proxy for environmental problem pressure that is otherwise difficult to grasp for environmental policy in general. The influence of green parties is measured with three indicators: electoral success, membership in parliament and participation in government. The scores are calculated as follows: a score of 1 (instead of 0) is assigned if the national vote share of green parties is above 1 per cent, considering the closest national election for the lower chamber to the four points in time. An additional score of 1 is assigned if green parties are represented in parliaments. Finally, a country receives a third score of 1 if green parties govern at the time of measurement or if the government representation has not ended more than one year before. Thus, the overall score of the variable for political demand similarity varies between 0 and 3 (see table A4.3 in the annex).

Two further variables express environmental problem pressure: population density per capita/km² and the level of environmental pollution as measured by CO₂ emissions per capita (see table A4.3 in the annex). Finally, the effects of pre-existing similarity of policies on convergence in later periods will be investigated in the pair approach. An overview over all variables and indicators is given in figure 4.2.

For most of the above variables, well-known and widely used data sources exist. Other variable data are simple proxies for the concepts, e.g., common border or common language, expressing geographical proximity and cultural similarity, respectively. For prevalent religion, we consulted and compared three different data sources. Only the data on green party strength, expressing political pressure, were collected from a wider range of published and often used sources as well as from existing data sets. Table A4.5 summarises the other variables and the data sources that were consulted.

<i>Dependent variable</i>	<i>Main independent variables</i>	<i>Other variables</i>	<i>Indicators</i>
Policy convergence Policy similarity	Economic interlinkage		Trade openness Bilateral trade openness
	Institutional interlinkage		
	Obligatory potential		Institutional membership, weighed by – Encompassingness – Type of law – Monitoring capacity
	Communicative potential		Institutional membership, weighed by – Encompassingness – Length of membership – Frequency of interaction – Permanence of delegation – Number of organisational bodies – Size of staff
		Cultural similarity	Language, religion Geographical proximity
		Environmental problem pressure	CO ₂ emissions Population density
		Level of income	GDP per capita
		Political demand	Green parties
		Pre-existing similarity	Policy similarity in t–1

Figure 4.2 Variables and indicators

4.4 CONCLUSION

This chapter provides an overview of the research design of this study and of the operationalisation of the variables. As to the dependent variable, environmental policy convergence, there are two important innovations. First, environmental policy convergence is based on a very broad selection of individual measures, namely forty policies representing several policy types and covering the full range of environmental media, including metrical data on limit values. Second, as the study strives to test theories that have a micro foundation and focus on governmental decisions, we use policy output data rather than outcome data. These data had to be collected for twenty-four countries and a period of thirty years. The process of data collection, in particular for the environmental policies but also for the institutional data, was very demanding and almost a 'project in its own right'. As to the main independent variables, the construction of the institutional indicators deserves attention. This project constitutes the first attempt to measure the integration of countries in international institutions in order to estimate their potential to harmonise policies and their potential to enhance policy learning among the member states.