

# **Explaining compliance: What makes the polluter turn green?**

**An event history analysis of member state compliance with the Montreal  
Protocol**



**Åsta Dyrnes Nordø**

**Master thesis**

**DEPARTMENT OF COMPARATIVE POLITICS**

**UNIVERSITY OF BERGEN**

**September 2010**

## **Abstract**

This thesis seeks to explore which structural factors that move member states into compliance with International Environmental Agreements (IEAs). The research question is: “under what conditions do members to an international environmental agreement comply with their commitments?”

Previous conducted studies on environmental accords have focused mainly on institutional treaty characteristics when explaining variance in compliance. The interest in regime factors has led to a focus on overall compliance more than on understanding member state compliance records. This lack of attention to the importance of extralegal factors to member state compliance with environmental agreements has been a motivation for this thesis.

As this thesis is exploratory by nature, the theoretical framework takes a broad approach, introducing country-specific determinants that are hypothesised to influence a state’s environmental record. To test the hypotheses, an event history analysis of signatory state compliance with the Montreal Protocol is conducted. 154 member states’ event histories between 1989 and 2008 are analysed and compared. The results show that as a signatory state’s length of protocol membership increases, regional compliance records improve and its general welfare rises, the associate has significantly better chances of experiencing compliance with the Montreal Protocol than a member state not inhabiting such features. Nonetheless, a powerful member state is significantly less likely to adhere to the Montreal accord obligations. From these findings, the effect of regional compliance is especially intriguing as it is poorly theorised in the literature. Furthermore, the analysis shows that several of the most theorised covariates fail to affect the propensity for compliance on the Montreal Protocol. Most surprising is the finding that neither political regime nor the scope of active ENGOs affects compliance rates significantly.

The results from this analysis indicate that one can not fully understand the variation in compliance with the Montreal Protocol without taking factors external to the environmental accords into consideration.

## Acknowledgements

I owe my thanks to many people who have provided me with invaluable help in the process of writing this thesis.

Thanks to my supervisor, Kristin Strømsnes, for excellent guidance and assistance throughout the process of writing this thesis. Thank you for your flexibility, your thorough feedbacks and for sincerely caring about me. Even though I have been a loner, you have done your best to make sure that the best guidance possible was given. For that I am grateful.

Also, my thanks to Michael Alvarez and the Methods group at the Department of Comparative Politics for great feedback and guidance at a crucial point of my thesis. Moreover, I owe a debt of gratitude to Espen for priceless help when the equations became tough.

Sincere thanks to Professor Eric Neymayer, London School of Economics and Political Science, Professor Ronald Mitchell, University of Oregon and Professor Helmut Breitmeier, University of Hagen, for taking your time to answer my e-mails and encouraging me.

Ullaloftet and the people there also deserve a big thank you, Rebekka, Kirsti, Terje and Ole, for all the late nights spent together. Also thanks to Asle and Kirsti for commenting on different chapters. Special thanks to Martin and Ingrid, for invaluable help as well as pep talks and free dinners along the way.

A big thank you to my family, for always being there.

Finally, a last thank you to Ole, you are my star.

Åsta Dyrnes Nordø

Bergen, October 2<sup>nd</sup> 2010

# Table of contents

<b>1. INTRODUCTION.....</b>	<b>0</b>
1.1 THEORETICAL RELEVANCE AND SCIENTIFIC CONTRIBUTION.....	1
1.2 STRUCTURE OF THE THESIS.....	3
<b>2. THE DEVELOPMENT OF IEAS AND THE MONTREAL PROTOCOL.....</b>	<b>4</b>
2.1 THE DEVELOPMENT OF INTERNATIONAL ENVIRONMENTAL REGIMES.....	4
2.2 THE MONTREAL PROTOCOL ON SUBSTANCES THAT DEplete THE OZONE LAYER.....	7
<b>3. THEORETICAL FRAMEWORK.....</b>	<b>9</b>
3.1 DEFINING COMPLIANCE.....	9
3.1.1 IMPLEMENTATION, COMPLIANCE, EFFECTIVENESS.....	10
3.2 UNDERSTANDING COMPLIANCE – THE POINT OF DEPARTURE.....	11
3.3 THE ENFORCEMENT SCHOOL.....	13
3.3.1 STATE POWER.....	13
3.3.2 ECONOMIC INTEGRATION.....	14
3.4 THE MANAGERIAL SCHOOL.....	15
3.4.1 GOVERNMENT AUTONOMY AND CAPACITY.....	15
3.4.2 POLITICAL REGIME.....	17
3.4.3 ECONOMIC DEVELOPMENT.....	19
3.4.4 LEGAL ORIGIN.....	20
3.5 THE LEGITIMACY SCHOOL.....	21
3.5.1 PUBLIC OPINION.....	21
3.5.2 MONTREAL PROTOCOL MEMBERSHIP LENGTH.....	22
3.5.3 EU MEMBERSHIP LENGTH.....	23
3.5.4 ENVIRONMENTAL NON-GOVERNMENTAL ORGANISATIONS.....	24
3.5.5 POLITICAL ORIENTATION.....	24
3.5.6 REGIONAL COMPLIANCE.....	25
3.6 SUMMARY OF ARGUMENTS AND HYPOTHESES.....	26
<b>4.0 METHODOLOGY: EVENT HISTORY ANALYSIS.....</b>	<b>28</b>
4.1 CHOOSING A QUANTITATIVE MODEL.....	28
4.2 CASE SELECTION – CHOOSING THE MONTREAL PROTOCOL.....	30
4.3 EVENT HISTORY ANALYSIS.....	31
4.3.1 DEFENDING THE CHOICE OF MODEL.....	33
4.3.2 INTRODUCING FUNDAMENTAL CONCEPTS: SURVIVAL AND RISK.....	34
4.3.3 THE UNDERLYING LOGIC OF EVENT HISTORY MODELLING.....	35
4.3.4 INTERPRETING THE HAZARD RATE.....	37
4.4 APPLYING THE EVENT HISTORY MODEL TO THE ANALYSIS OF MEMBER STATE COMPLIANCE WITH THE MONTREAL PROTOCOL.....	38
4.4.1 CHOOSING EVENT HISTORY MODEL.....	38
4.4.2 THE COX PROPORTIONAL HAZARDS MODEL.....	39
4.4.3 MODELLING GROUP EFFECTS: THE STRATIFIED COX MODEL.....	40
4.5 METHOD-SPECIFIC ISSUES: PROPORTIONALITY AND TIME-VARYING COVARIATES.....	42
4.5.1 TESTING THE PROPORTIONALITY ASSUMPTION.....	42
4.5.2 CAPTURING THE FUNCTIONAL FORM OF TIME-VARYING COVARIATES.....	43

<b>4.6 ANALYSIS SPECIFIC CONSIDERATIONS .....</b>	<b>46</b>
<b><u>5.0 DATA AND VARIABLES.....</u></b>	<b><u>48</u></b>
<b>5.1 DATA .....</b>	<b>48</b>
5.1.1 THE DATASET.....	48
5.1.2 MODEL-SPECIFIC ISSUES: CENSORING AND TRUNCATION .....	49
5.1.3 SAMPLE SELECTION .....	50
<b>5.2 THE DEPENDENT VARIABLE: TIME UNTIL COMPLIANCE WITH THE MP .....</b>	<b>51</b>
<b>5.3 OPERATIONALISATION OF THE INDEPENDENT VARIABLES .....</b>	<b>54</b>
5.3.1 ENFORCEMENT FACTORS .....	55
5.3.2 MANAGERIAL FACTORS .....	55
5.3.3 LEGITIMACY FACTORS .....	60
<b><u>6. ANALYSING THE DETERMINANTS OF COMPLIANCE WITH THE MONTREAL PROTOCOL.....</u></b>	<b><u>64</u></b>
<b>6.1 DESCRIPTIVE STATISTICS – RANGE AND DISTRIBUTION .....</b>	<b>64</b>
6.1.1 THE DEPENDENT VARIABLE .....	65
6.1.2 ENFORCEMENT DETERMINANTS .....	66
6.1.3 MANAGERIAL DETERMINANTS .....	66
6.1.4 LEGITIMACY DETERMINANTS .....	67
<b>6.2 BIVARIATE COX REGRESSIONS .....</b>	<b>68</b>
<b>6.3 MULTIVARIATE COX REGRESSIONS.....</b>	<b>70</b>
6.3.1 COMPLIANCE WITH THE MONTREAL PROTOCOL – THE RESULTS .....	71
6.3.2 MODEL ONE .....	74
6.3.3 MODEL TWO.....	76
6.3.4 MODEL THREE .....	77
<b>6.4 CONSTRUCTING A SYNTHESIS MODEL .....</b>	<b>78</b>
<b>6.5 DISCUSSION.....</b>	<b>82</b>
<b>6.6 SUMMARY OF THE FINDINGS .....</b>	<b>89</b>
<b><u>7. CONCLUSION.....</u></b>	<b><u>92</u></b>
<b>7.1 EXPLAINING COMPLIANCE WITH THE MONTREAL PROTOCOL.....</b>	<b>92</b>
<b>7.2 SUGGESTIONS FOR FURTHER RESEARCH.....</b>	<b>93</b>

## List of Tables

Table 1: Expected effects on member state compliance with the Montreal Protocol.....	26
Table 2: Overview of the dataset.....	49
Table 3: Descriptive statistics .....	65
Table 4: Bivariate Cox regression of compliance with the Montreal Protocol.....	68
Table 5: Multivariate analysis of the Montreal Protocol.....	73
Table 6: Overview of variables with significant effects in Models 1, 2 and 3.....	79
Table 7: Synthesis models.....	80
Table 8: Wald test of the Green popularity and EU-length covariates' joint effect.....	81
Table 9: The findings of the analysis .....	90
Table 10: Countries included in the dataset .....	103
Table 11: Codebook .....	104
Table 12: Schoenfeld tests, run on all models.....	105
Table 13: Pairwise correlations .....	108
Table 14: Reported hazard ratios from Green party popularity at various points in time.....	109
Table 15: Change in hazard ratio from a one million increase in population .....	110
Table 16: Change in hazard ratio from one additional year as member of the Montreal Protocol .....	111

## List of Figures

Figure 1: Cumulative counts of international environmental activities, 1870-1990.....	4
Figure 2: Baseline hazards for the two groups of Montreal Protocol members.....	106
Figure 3: Cox-Snell residuals, run on the biggest model (Model 1).....	106
Figure 4: The shape of the hazard for the different models, based on Table 15 .....	110
Figure 5: The shape of the hazard for the different models, based on Table 16.....	111

## **Abbreviations**

CFC = Chlorofluorocarbons

CITES = the Washington Convention on International Trade in Endangered Species of Wild Fauna and Flora

CPI = Corruption Perceptions Index

ENGO = Environmental Non-Governmental Organisation

EU = European Union

EVS = European Values Survey

GDP = Gross Domestic Product

HDI = Human Development Index

IEA = International Environmental Agreement

ILO = International Labour Organisation

IMF = International Monetary Fund

IR = International Relations

MP = Montreal Protocol

NO<sub>x</sub> = Nitrogen Oxide gases

ODP = Ozone Depleting Potential

ODS = Ozone Depleting Substances

SO<sub>2</sub> = Sulphur dioxide

UNEP= United Nations Environmental Programme

WDI = Worldwide Governance Indicators

WVS = World Values Survey





“The political science literature to date has focused on how regimes influence the environmental behaviours of states, but it could benefit by framing the question as what explains variation in the environmental behaviour of states? This subtle shift directs our attention to the many nonlegal drivers of environmental behaviours that are often arrayed against international environmental agreements but sometimes facilitate their efforts (Mitchell 2003:449).

“Questions of compliance – to what extent do states comply, which states are likely to comply, what patterns of compliance exist within and across areas of regulation? – have not been extensively investigated and remain poorly understood” (Haas 1998:17).

## 1. Introduction

The last century has been characterised by a pronounced increase in international regimes of cooperation. Since the 1970s, a growing concern with the impact industrial progress and human activity has on the environment has led to the establishment of a series of International Environmental Agreements (IEAs). One of the most well-known traits of supranational regimes is that they do not impede state sovereignty. As IEAs lack the efficient regulator means that a state inhabits, international agreements do not have the power to force members to act. Why, then, do states follow international environmental law? This puzzle has provoked this thesis’ research question; *under what conditions do members to an international environmental agreement comply with their commitments?*

Environmental issues are typical commons issues. As states cannot seal their borders from environmental degradation, and many problems affect what international law considers commons areas, states cannot privatize economic or social damages resulting from pollution, deforestation or depletion of the ozone layer. Hence, international collaboration is necessitated to enable changes. Nonetheless, the development of IEAs is a recent phenomenon, and their success at changing state behaviour is heavily discussed (see for example: Chayes and Chayes 1993; 1995; Checkel 2001; Downs et al. 1996; Haas et al. 1993; Mitchell 1994; Simmons 1998). In my contribution to this discussion, I focus on member state compliance. *Compliance* is a notion reflecting behavioural change among members to an agreement in legal terms. Exactly this thesis’ interest in actual change is what lies behind the choice of focus<sup>1</sup>.

---

<sup>1</sup> A more thorough definition and discussion of the concept is introduced in section 3.1.

When describing differences in IEA outcomes, scholars within the International Relations (IR) school tend to focus on the framework of agreement in their analyses, treating nation-specific indicators as constants. The use of enforcement and management mechanisms has been thoroughly covered. However, several influential studies done on this field end up discussing factors external to the agreement, and how they might matter (Brown Weiss and Jacobson 1998; Cameron et al. 1996; Haas 1998; Mitchell 2003). It is from such discussions my thesis departs.

In the process of answering my research question, the *Montreal Protocol on Substances that Deplete the Ozone Layer* (henceforth referred to as the Montreal Protocol) was chosen as the case under scrutiny. To be able to study member states' behaviour to the Protocol, a statistical analysis labelled event history analysis is applied. Three different models are run, followed by the construction of a final model visualising the variables most successful at explaining the variance in associate compliance seen in the Montreal accord.

### **1.1 Theoretical relevance and scientific contribution**

International Relations theorists have emphasized the importance of the regimes'<sup>2</sup> institutional design when considering state compliance. This is connected with a scholarly focus on responses to non-compliance (coercion, institutional compromise and social learning) rather than understanding which mechanisms drive compliance (Haas 2008:xvii). Consequently, most of the literature on the mixed success of environmental governance has been considered an institutional problem rather than a structural one. However, recent empirical research point to the importance of factors external to the treaty, such as strong civil engagement, democracy and effectiveness of the bureaucracy (Bernhagen 2008; Binder and Neumayer 2005; Neumayer 2002, 2003; Roberts et al. 2004; Simmons 2000; Perkins and Neumayer 2007). I find this partition intriguing, and wish to search deeper for an understanding of which extralegal mechanisms influence the chance of reaching compliance. Thus, this thesis is an attempt to shed new light on what explains compliance based on an approach focusing on structures (like government autonomy, political regime and legal origin) and processes (like economic development and integration in the world economy) characterising the member states.

---

<sup>2</sup> This thesis will only use the designation 'regime' in the context international environmental regime, and not in the sense of a political system or synonymous with state. This is done to avoid confusion.

Moving the focus from treaty-induced explanatory variables to more structural covariates, the natural choice of units of analysis also changes. In most of the literature on IEA compliance and effectiveness, researchers are preoccupied with overall compliance. Thus, little attention has been given to state-specific behaviour within the framework of IEAs, and if this focus has been present, it has been restricted to a focus on behaviour among the most powerful states (see for example: Breitmeier et al. 2006). Hence, I argue that my thesis brings new insights to the literature by shifting the focus to member states and examining the processes driving compliance and defiance at the national level.

The original goal of this thesis was to compare two international environmental regimes and see whether the same determinants played a significant role in making member states comply. Two datasets were created, but because of the limited time and scope of a master thesis, as well as methodological issues connected with the second dataset, I decided to analyse determinants' effect on compliance with the Montreal Protocol alone. This prevents me from comparing the importance of different determinants between IEAs. However, as Mitchell (2004:128-131) argues, comparability is questionable in the case of international agreement, because of the heterogeneity of existing agreements as well as in regime goals and in the difficulty of inducing actual change. Thus, a motivation is to point out which determinants improve the chances of a country moving into compliance with the Montreal Protocol. This will provide increased knowledge about which structural factors play a role in such complex 'two-level' processes linking international and national politics (Putnam 1988).

There are serious obstacles presented when wishing to evaluate the outcome of international environmental accords. First, most agreements are so recent that estimating their effects are premature. Secondly, older protocols may inhabit important information, but the data are often neither consistent nor reliable and trustworthy results are thus hard to obtain (Ringquist and Kostadinova 2005). Thirdly, it is hard to decide when a state has changed its behaviour sufficiently to be labelled a complier. Following Brown Weiss and Jacobson (1998:2): "(...) compliance occurs across a scale of shades of grey". Such challenges have made qualitative approaches the dominant method when conducting research on the outcome of IEAs. Thus, when I apply a statistical method to answer the research question, this should represent a valuable contribution to the predominantly qualitative literature characterising the paradigm. As such, this study has an exploratory nature.

## **1.2 Structure of the thesis**

The Montreal Protocol making up the framework for analysis is not familiar to everyone. Thus, *chapter two* discusses the historical development of IEAs and gives an introduction to the Montreal Protocol. Then, *chapter three* presents the theoretical framework for the thesis. Here, different theories are presented and made measureable through clearly stated hypotheses. Thereafter, *chapter four* introduces event history analysis, which is the methodological framework chosen for this thesis. Furthermore, in *chapter five* the data material is discussed and the operationalisation of hypothesis into variables is introduced. *Chapter six* reports the results that can be inferred from the statistical analysis divided into three models. Based on the findings here, a synthesis model is established to summarize the statistically significant results. Finishing off, the regression results are discussed in light of the hypotheses set out in the theory chapter. Finally, *chapter seven* concludes the research question by summarizing the findings and discussing the implications of this analysis.

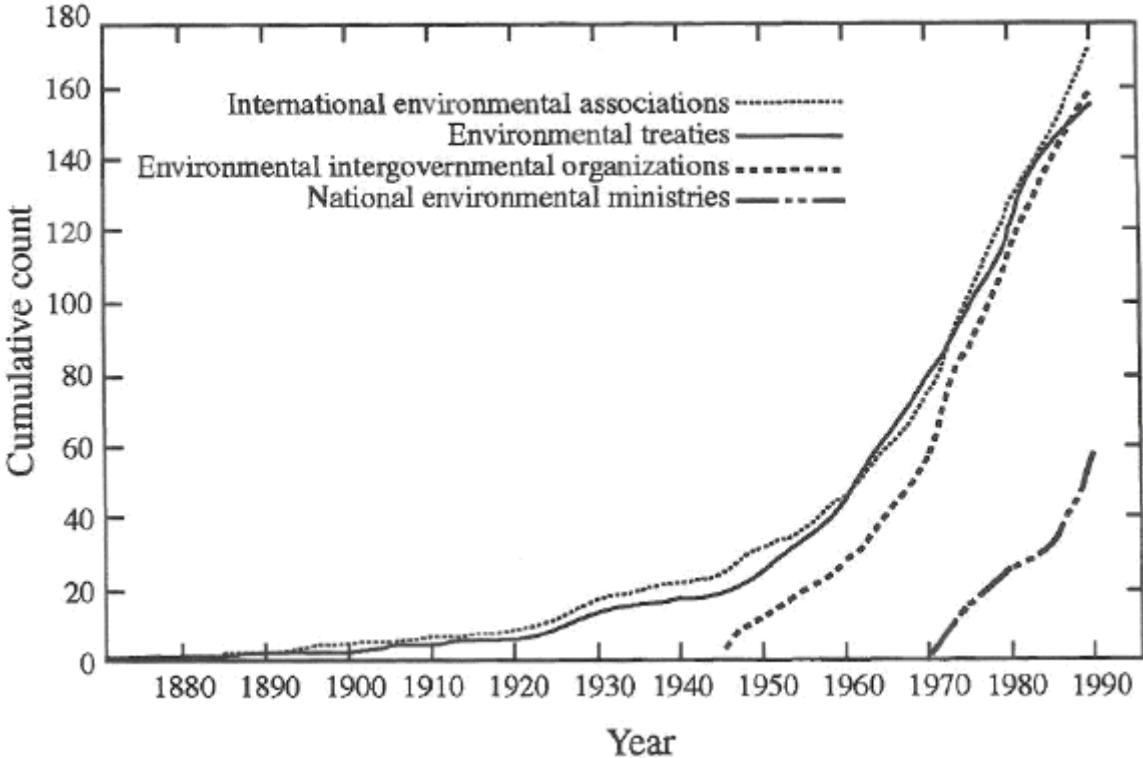
## 2. The development of IEAs and the Montreal Protocol

After the Second World War, an explosion in mutual commitment through international agreements has occurred. Cross-national green agreements, focusing on the environmental damage seen as the industrial West advanced, had been present from the onset of the 20<sup>th</sup> century, but not to the extent and comprehension witnessed from 1970 and onwards. Today, there exist approximately 225 active international environmental accords in the world (Barrett 2003). This chapter first includes a review of the development and structuring of international environmental regimes. Then, the history of the IEA under scrutiny in this thesis, the Montreal Protocol on Substances that Deplete the Ozone Layer<sup>3</sup> is presented and discussed.

### 2.1 The development of International Environmental Regimes

In this thesis, an international environmental regime is understood as “one that is primarily focused on some aspect of the relationship between human society and nature and that involves three or more nation state parties” (Meyer et al. 1997:636). As displayed in Figure 1, the establishment of environmental commitment across states is a rather new phenomenon.

Figure 1: Cumulative counts of international environmental activities, 1870-1990



The figure is taken from Meyer et al. (1997:625)

<sup>3</sup> A thorough treatment of case selection and operationalisations is introduced in chapters 4 and 5.

The first international agreements were formed between European states and were mainly motivated by one of two objectives. Either, an understanding of nature as a resource for allocation was behind mutual commitments. For instance, the first international environmental convention was signed in 1900 and included protection of animals, birds and fish in Africa. Of course, the treaty was initiated by the colonial powers who wanted to secure the stock of wild animals for their own hunters. Or, some agreements were based on a sentimental view on some aspects of nature, like the 1933 Convention Relative to the Preservation of Fauna and Flora in their Natural State (Andresen et al. 2008:22-23; Meyer et al. 1997). In the period up until World War II few new IEAs were established. Nonetheless, this period saw increased nongovernmental associational activity, although in a decentralised vein.

A crucial happening for the modern IEAs was the establishment of the United Nations (UN) in 1945. Suddenly, there was an arena for questions of an international character, like human rights, security and defence, trade and environment. Furthermore, the structural frame of the UN served as an inspiration for the institutionalisation of environmental commitment. Despite such advances, the bulk of international treaties seen in the 1950s were related to security and trade issues, and environmental damage was given little attention. This improved in the 1960s as an increased public involvement in environmental issues culminated in the 1972 UN Conference on the Human Environment (the Stockholm Conference), where 120 states participated. The conference was a breakthrough for environmental cooperation with the establishment of the United Nations Environmental Programme (UNEP) as well as important principles with regard to environment and development. UNEP is today the single most important world organisation concerned with the environment, and is a secretariat to several of the biggest IEAs (Andresen et al. 2008:23-25).

In the wake of the conference, a sharp increase in IEAs were seen, quite a few of them becoming global in their reach. Examples are Convention on Long-range Transboundary Air Pollution (CLRTAP) (1979), United Nations Convention on the Law of the Sea (1982), the Vienna Convention for the Protection of the Ozone Layer (1985) and the International Tropical Timber Organisation (1986) (Andresen et al. 2008). Nevertheless, nature did not fully enter the international agenda until the discovery of the ‘ozone hole’ over the Antarctic and the devastating Chernobyl disaster in the mid-1980s. This coincided with the UN report, *Our Common Future*, which established the notion sustainable development<sup>4</sup> and concluded

---

<sup>4</sup> The meaning behind the term is that nature and development needs to be jointly considered (Andresen et al. 2008:24).

that ‘the polluter pays,’ leaving the main responsibility for environmental cleanup with the developed countries. This acknowledgement made the developing countries involve themselves to a much greater degree than what had been previously seen. The last twenty years has not seen too many new IEAs as new environmental challenges are included into the structures of the already existing conventions<sup>5</sup>. Despite few new IEAs being seen, overall international environmental activity has increased substantially in this period (Meyer et al. 1997:638).

What explains the sudden explosion in environmental commitment after 1960s, as visualised in Figure 1? Environmental degradation has been occurring for several centuries without collective action being initiated. Examples are the fishery collapses seen during the period of Hanseatic League or the quadrupling of crop land from 1700 to 1920 leading to enormous losses to biodiversity (Meyer et al. 1997:637-638). Meyer et al. (1997:629-639) argue that there are two dramatic changes in world society explaining the rise of contemporary environmental regimes. First, the increased prominence of scientific analyses of the nature that was able to define and codify environmental degradation and communicate these findings to a global audience, thus creating collective mobilisation and action. As the massive expansion seen in scientific activity established a notion of an interdependent ecosystem which sustains the very possibility of life, a frame for international activity was created. Secondly, the creation of a broad world organisational structure, most importantly through the United Nations, was important because it provided an international arena that encouraged mobilisation around issues transcending state borders.

To sum up, Meyer et al. (1997:637) gives a good description of the change seen in IEAs’ structure and focus over time.

“Early treaties tended to be specific, signed by limited numbers of developed countries, and concerned with the management of specific international dependencies” whereas “(...) more recent treaties (...) emphasize regional and global interdependencies, and they are rooted in a broad and universalistic scientific conception of nature as an ecosystem with which human society must come into balance”.

It is the latter type of agreement described here this thesis will work with.

---

<sup>5</sup> Notable exceptions are the UN Framework Convention on Climate Change (UNFCCC) (1992) and the linked Kyoto Protocol as well as Convention on Biological Diversity (CBD) (1992).

## ***2.2 The Montreal Protocol on Substances that Deplete the Ozone Layer.***

The observation of a continued weakening of the ozone layer, and especially the discovery of the enormous ‘ozone hole’ over the Antarctic in 1985, led to renewed attention to the causes and effects of the depletion of the ozone layer. Since the 1970s, uncertainty had existed with regard to how severe the effect of Chlorofluorocarbon (CFC) and halon gases was on the ozone layer. Despite the fact that the chemists who discovered the destructive effect of CFC gases on the ozone layer received the Nobel Prize in chemistry for their work, powerful chemical companies worked against this recognition, dismissing it as mere science fiction (Andresen et al. 2008:56). However, with the disclosure of the Antarctic ‘ozone hole,’ international scientific agreed that the problem was real and necessary to counteract. Reducing CFC and halon gas production was vital in combating the reduction of the earth’s ozone layer. Yet, this was seen as a great challenge, as use of both CFC and halon gases had become widespread as important components in refrigerators and freezers, spray cans, cellular plastics, solvents, dry cleaners and fire-extinguishers (Andresen et al. 2008: 56-60).

Previous to the groundbreaking findings, international cooperation had been initiated. Notwithstanding, as a reflection of the uncertainty and the market situation with Europe and North America being the largest consumers and producers, the first international agreement, named the Vienna Convention for the Protection of the Ozone Layer, did not oblige the countries to specific reductions. Thus, the negotiations of a new protocol to the Vienna convention were tough, with the EU being split in two between the countries who followed the US in their fight for a more radical goal of stabilisation in the short run and full stop in the long run and those favouring more moderate actions. The negotiation round ended with the EU giving in to the demands of the more ambitious side. Thus, 24 countries and the EU agreed on the Montreal Protocol in 1987. It was ratified in 1988 and entered into force on January 1<sup>st</sup> 1989, as a supplement to the Vienna convention.

The Protocol objective was a 50 percent reduction of the production and consumption of CFCs and halons in the industrialised member states by 1999, the baseline year being 1986. UNEP served as the secretariat to the agreement. Developing countries were given a ten year postponement to reach the same goal (referred to as Article 5- members in this thesis)<sup>6</sup>.

---

<sup>6</sup> According to the Protocol text: “Any Party that is a developing country and whose annual calculated level of consumption of the controlled substances is less than 0,3 kilograms per capita on the date of the entry into force of the Protocol for it, or any time thereafter within ten years of the date of entry into force of the Protocol shall, in order to meet its basic domestic needs, be entitled to delay its compliance with the control measures set out in paragraphs 1 to 4 of Article 2 by ten years after that specified in those paragraphs” (Benedick 1991:235-236).



Despite this, pulls towards an even more radical regime ended in changes in the original objective of the non-article 5 members. Fifteen months after the Montreal Protocol had been agreed on, the US and the EU urged for introducing a stricter regime requiring full stop in the production and use of CFC and halon gases by January 1<sup>st</sup> 2000. The more ambitious plan was formally accepted by all parties on the second meeting of the Parties in London in 1990. Later on, the non-article 5 group decided on even more radical phase-out plans, pushing the deadline forward to 1996 and 1994 for CFCs and halons respectively. At this point in time, there were 58 Parties to the Montreal Protocol, making up 90% of the world's production and consumption of ozone reducing substances. As of November 26<sup>th</sup> 2009, 196 countries had ratified the Montreal Protocol (The Ozone secretariat 2010).

New amendments were added to the Convention, expanding the area of responsibility and enhancing deeper cooperation on the issue of Ozone Depleting Substances (ODS). The London Amendment added the Annex B substances carbon tetrachloride, methyl chloroform and other halogenated CFCs to the Protocol. The Copenhagen Amendment added the Annex C substances hydrochlorofluorocarbons (HCFCs), and hydrobromofluorocarbons (HBFCs), as well as the Annex E substance methyl bromide to the original treaty (Andresen et al. 2008:61). This thesis does not consider these amendments any further.

Among scientists working on the field of IEAs, it is well known that the Montreal Protocol is described as the success story in the (rather short) history of international environmental cooperation. Thus, choosing this as my case, I am aware that this might strike the reader as both biased and not thoroughly thought through. Nonetheless, my focus is not on whether member states comply, but rather on *understanding which factors drive compliance with IEAs*. This difference, although subtle, is crucial to the thesis and defends the choice of the Montreal Protocol as the case of departure.

### 3. Theoretical framework

Compliance with the Montreal Protocol is, strictly defined, a legal question. Still, this thesis argues that the process towards compliance is too complex to be properly understood without taking extralegal features inherent to each state into account. Policy-makers are confined by economical, political, social and cultural structures in society, as much as by the public will and their own personal convictions. This chapter will discuss the definition of compliance, and implement it in a broader clarifying discussion of the important distinction between compliance, implementation and effectiveness in the international regime literature.

Furthermore, theories and previous studies are included in a discussion making up the background for the hypotheses that this thesis aims to test empirically. To structure this examination, three prominent compliance approaches in the IR literature have been used as a frame of reference.

#### 3.1 Defining compliance

The term compliance is commonly used when comparing behaviour to specific treaty provisions, a treaty's broader spirit and principles, implicit international norms and informal agreements (Mitchell 1996:5). The different ways of understanding compliance reported here are manifested in different definitions presented by scholars. Chayes and Chayes (1993:4) operate with a rather abstract definition. They argue that compliance is successful when states: "(...) alter their behaviour, their relationships, and their expectations of one another over time in accordance with its terms." Young (1979:3) moves close to Chayes and Chayes' understanding in his classical work on compliance. According to his definition:

"compliance can be said to occur when the actual behaviour of a given subject conforms to prescribed behaviour, and non-compliance or violation occurs when actual behaviour departs significantly from prescribed behaviour."

These definitions offer a broad understanding of compliance, not only confining it to *de jure* compliance, but to an internalisation of the normative basis for international environmental cooperation. Following this line of argument, perfect compliance is exceedingly hard to achieve and even harder to measure. Therefore, these scholars further argue that substantial compliance is what should be sought by the treaty makers. The judgment of what substantial

compliance actually signifies is left to the researcher, allowing considerable subjectivity into an analysis.

A more tangible definition is offered by Brown Weiss and Jacobson (1998:4), who define compliers as the member states who: “adhere to the provisions of the accord and to implementing measures that they have instituted.” Mitchell (1996:5) acknowledges that treaties may induce positive change in behaviour among the actors involved, but chooses to exclude the notion of compliance with the norms underlying the agreement. He thus offers an even more narrow definition where compliance is regarded as: “an actor’s behaviour that conforms to a treaty’s explicit rules.”

All four definitions introduced share a focus on subject behaviour. This adds an important qualification to the compliance issue, demonstrating that when doing research on compliance, *actual behaviour of subjects* needs to be the focus, not intentions and attitudes. This means that intentions shown through membership and positive attitudes towards an IEA does not help move a country into compliance unless action is taken at the national level through policy changes that materialise in actual changes in society.

The definitions presented above show a lack of consensus among scholars with regard to what should be included in the concept *compliance*. It is natural that such an understanding varies with how the definition is used in a study. Chayes and Chayes (1993:176) argue that the treaty regime should be measured as a level of overall compliance, considering the treaty’s scope and goals. Mitchell (1996:6), on the other hand, argues that measuring overall compliance with an accord is likely to lead to useless inferences, calling for an empirical measuring of compliance directed against treaty provisions. As this thesis is based on a statistical approach, two things are particularly important; first, that the concepts utilized are not biased and second, that they can be further operationalised into tangible measures susceptible to empirical testing. Based on this, it naturally follows that the narrow definitions presented by Brown Weiss and Jacobson and Mitchell are most suitable for my analysis.

### **3.1.1. Implementation, compliance, effectiveness**

The term compliance is used somewhat inconsistently in the existing literature on IEAs, making it easy to confuse with related terminology such as implementation and effectiveness. To avoid confusion, I carefully distinguish these terms from each other as they refer to different aspects of the political and legal cooperation process (Axelrod et al. 2005: 163-164)

Following Young (1979) and Brown Weiss and Jacobson (1998:4-6), I differentiate between implementation and compliance. *Implementation* refers to the measures that governments take to translate international accords into domestic law and policy. The focus here is on the procedures decided upon by the national authorities to secure compliance further along the line. *Compliance*, then, goes beyond implementation, and is harder to measure. As mentioned in the previous section, compliance is a matter of whether and to what extent member states concur to the provision of the agreement. Thus, domestic policy must translate into action that helps reaching the stated goals of the agreement. To illustrate the difference, one can say that a member state that is successful at implementing the Montreal Protocol is a state that effectively incorporates the laws put forward by the accord into the national legal system. Yet, to be a good complier to the Montreal Protocol, this legal system must be enforced on all CFC and halon polluters, and actual action must be taken to reduce the emission levels to the level agreed upon, within the deadline set by the Ozone Secretariat.

Furthermore, it is important to clarify the distinction between the terms *compliance* and *effectiveness*. Both concepts are considered to evaluate the outcome of a process. The difference then, is what they emphasise when evaluating a regime. As Underdal (1998:6) points out, the level of compliance, together with the stringency of regime injunctions and side effects, make up the functions that together can work as a measure on effectiveness. Accordingly, level of compliance can partly explain level of effectiveness. Compliance considers the legal aspect of the outcome; whether the member states adhere to the rules it has bound itself to adhere to. Effectiveness looks at the broader consequences, for example whether changes can be traced in the environment due to the policies initiated by an IEA. Such causal connections are hard to verify. Thus, as Brown Weiss and Jacobson (1998) underline, even though signatory states are in compliance with a treaty, it does not mean that the treaty is effective in fighting the problems it addresses. Closely connected to this argument, Underdal (1998:6) asserts that;

“(...) maximizing compliance is not necessarily a good design principle if the ultimate goal is to enhance regime effectiveness. The regimes that are most easily complied with are those that require little or no change of behaviour.”

### **3.2 Understanding compliance – the point of departure**

The compliance debate first became salient in the IR literature in the 1970s. It derived from the literature on international arms control which was the main field of research during the

Cold War. Following the fall of the Iron Curtain, the focus on compliance was transferred into other areas of international cooperation. After a few years with an extensive literature being produced on IEAs and compliance in the first half of the 1990s, Mitchell (1996:4) concluded that: “the number of and variety of proposals to improve international environmental treaties suggest (...) that we still lack a solid understanding of what factors facilitate, and which impede, compliance with a treaty”. Today, 20 years after the debate accelerated, the compliance literature is still factionalised.

The compliance paradigm originated as a debate between the realist enforcement school emphasizing the importance of sanctions, and the managerial school focusing on the importance of capacity building for securing compliance. Eventually, the constructivists also included themselves in the discussion of what makes states comply, introducing the importance of norms and legitimacy into the discourse. As Börzel et al. (2007) state, these theories have traditionally been preoccupied with evaluating regime framework as it has been used for research on the differences in level of compliance *between* international agreements. Such an approach is inappropriate in this thesis as I focus on *cross-national patterns of compliance with a single IEA*. Nevertheless, inspired by Börzel et al. (2007) and Bernhagen (2008), I argue that the IR theories projecting the field of environmental governance can be used as a framework for hypotheses on cross-national differences of compliance. Hence, this chapter is structured around the realist enforcement school, the institutionalist managerial school and the constructivist legitimacy school. A short introduction to each of the schools is presented, consecutively followed by the presentation of the theories and hypothesis that I find to fit under each branch. However, it needs to be pointed out that the intention of structuring the chapter this way is not a wish to test and compare the strength of each school up against the others. This is merely a structural grip to reflect the literature that the IEA paradigm originates from and place the hypotheses accordingly. Thus, this thesis is interested in assessing each hypothesis’ explanatory power when compliance is considered, and the use of IR theories should not be considered anything more than a structural device.

As already mentioned, this thesis is exploratory in nature. As the compliance paradigm with regard to cross-national variation is still under-studied, the theoretical foundation for the hypotheses to be tested here is less established than the theories that originated with regard to overall regime compliance and effectiveness. Consequently, the point of departure for the following theoretical discussion is theories and empirical assessments on the broader subject of international environmental politics. The aim of this study then, is to test a group of hypothesis that have been postulated to affect a state’s environmental record, and see whether

these assumptions hold in a quantitative analysis on member state compliance with the Montreal Protocol.

### **3.3 The enforcement school**

The enforcement school arose out of a realist account of international politics. Some realists argue that the only reason why states comply with IEAs is that they are so shallow and vaguely formulated that it is close to impossible for members to defect (Downs et al. 1996; Haas 1998). The enforcement school treats international accords in a more serious manner, but considers inter-state binding agreements to be signed by states only when the states' myopic self-interests justify the costs of membership. Non-compliance in the realist eyes, therefore, is intentional and based on a profit-maximising rationale. Consequently, the scholars contend that for members to comply with international agreements, institutionalised monitoring and sanctioning mechanisms that make the costs of defection exceed the benefits, must be introduced (Börzel et al. 2007; Downs et al. 1996). However, states do not face the same compliance costs, nor are they homogeneously sensitive to sanctions. This section will, through theories on state power and economic integration, consider the importance of autonomy versus interdependence when explaining the observed variance in compliance across member states to the Montreal Protocol.

#### **3.3.1 State power**

Following the realist line of argument, one can expect the sensitivity of imposed costs to vary with states' economic and political power (Börzel et al. 2007:5-8). Thus, powerful states are expected to be less preoccupied with complying with international accords as their size politically, economically or military is what grants them legitimacy and influence, and not their cooperation reputation (Perkins and Neumayer 2007:22). Along a similar line of argument, one can argue that powerful states are more autonomous, making their ability to resist pressure and defy international obligations stronger (Börzel et al. 2007:5-6). Hence, it naturally follows that external enforcement constraints are more costly for the economically and politically weak states. Based on this assumption, scholars expect that less powerful members ratifying IEAs are better compliers than their powerful adversaries. A competing assumption is that the more economic and political power a state has got, the more able it is to shape the legal framework according to its own preferences (Moravcsik 1997; Börzel et al.

2007). Thus, one should expect a better record of compliance to be seen among powerful states.

In their study of member state compliance with EU law, Börzel et al. (2007) find support for the theoretical assumption that less powerful states are better compliers than their adversaries. Yet, this stand is challenged by Neumayer (2002), who argues, based on a quantitative study of reporting requirements under CITES<sup>7</sup>, that one can expect big and important signatory states to be better compliers than smaller, less important members. His line of argument is that the process to reach compliance should not be taken as a sign of outright concern, but rather as an indicator of state's interest in demonstrating their importance in the world. The lack of consensus in the literature leads us to the following two hypotheses:

*H1: Compliance with the Montreal Protocol is more likely when the member state is weak.*

*H2: Compliance with the Montreal Protocol is more likely among strong member states.*

### **3.3.2 Economic integration**

The more dependent a state is on sustaining a good relationship with the outside world, the more it will be in its interest to comply with an agreement one is member of. Consequently, globally and regionally integrated states should be more prone to adopt ambitious agreements as well as change their behaviour accordingly compared to their less integrated adversaries (Mitchell 2003:453; Haas 1998:29-30). Based on empirical studies conducted on trade and environment by Runge (1996), Vogel (1997) and Péchoux and Pouyet (2002), Bernhagen summarises the relationship between environmental records and openness to trade as ambiguous.

“On the one hand, trade integration can increase environmental cooperation by promoting growth and technological development, thereby increasing public demands for environmental protection while reducing the pollution intensity of production. On the other hand, integration in global trade provides governments with incentives to keep the domestic cost burden on producers low, minimizing the costs to business from strong environmental regulation, and thereby depressing a country's commitment to international environmental action” (Bernhagen 2008:92).

As the first scholar to test this assumption quantitatively with regard to environmental compliance and participation, Bernhagen (2008) finds that trade dependence depresses IEA

---

<sup>7</sup> The Convention on International Trade in Endangered Species of Wild Fauna and Flora.

participation significantly, whereas no significant effect on compliance can be concluded. However, the small sample is problematic, possibly making the detection of statistically significant results unnaturally hard. The theoretical differences lead to the following two hypotheses:

*H3: Higher levels of integration in the world economy raise the propensity for compliance with the Montreal accord.*

*H4: Higher levels of integration in the world economy lower the likelihood of member compliance with the Montreal Protocol.*

### **3.4 The managerial school**

Opposed to the realist focus on a lack of willingness, the managerialists<sup>8</sup> focus on the states' *lack of ability* to comply with international accords. The managerial school is a branch within the institutionalist IR approach. The literature has focused on three main sources of involuntary defiance: weak or lacking state capacities, ambiguous and indeterminate treaty language and the temporal dimension tied to social and economic changes following treaty obligations (Chayes and Chayes 1993:188-197; Chayes and Chayes 1995; Haas et al. 1993; Brown Weiss and Jacobson 1998; Börzel et al. 2007; Breitmeier et al. 2006). Most of the empirical work on compliance has concluded that state capacity is decisive. This section will focus merely on state capacity and social and economic changes, as the treaty language apply similarly to all members of the accord, and therefore cannot account for cross-national variation. There are many factors that may influence state capacity, directly as well as indirectly. First, government capacity and autonomy is considered, followed by political regime, and economic development, which both make up a broader framework that restricts a state's freedom of movement in the international sphere. Lastly, the importance of legal origin is discussed as it frames the legal system through which international law is domesticated.

#### **3.4.1 Government autonomy and capacity**

Because the literature that considers state capacity uses the concept differently, an important divide between government autonomy and government capacity is introduced (Simmons 1998; Börzel et al. 2007). *Government autonomy* emphasises the dynamic between domestic institutional structures and partisan veto players. The underlying logic is that as the amount of institutional veto players increases, the state is less efficient in making and implementing

---

<sup>8</sup> The designation "managerial school" was first introduced by Downs et al. (1996).



national policies that comply with costly international prescriptions (Börzel et al. 2007:8; Haverland 2000; Giuliani 2003). As influential veto players come to the fore, the likelihood of a state moving into compliance decreases. Haverland (2000) finds support for the paralysing effect veto players can have in his case study of implementation of the Packaging Waste Directive imposed on the member states by the EU. He concludes that the timing and quality of compliance depends on the opposition met by institutional veto players when running the directive through the national political system. Similarly, Giuliani (2003:152) finds that the degree of national adaption to the EU decreases as the number of veto players increase, as they increase internal decision-making costs and creates less flexibility and effectiveness in the policy-making system.

*Government capacity* comprises a state's ability to act, based on financial endowments and human resources (Cameron et al. 1996:49). Still, sufficient endowments do not automatically mean full compliance. Dispersed resources may make coordination difficult, independent of the endowments available. Also, connected to state capacity is the existence of *corruption*. The more corruption is present in a society, the harder it is to secure that endowments are effectively managed. I thus hypothesise that a society driven through a high degree of corruption is more prone to defect than a society where corruption is insignificant.

The presentation above allows us to assume a positive relationship between government capacity and government autonomy and state compliance, as well as a negative effect of corruption on compliance. An influential study by Brown Weiss and Jacobson (1998) has compared state compliance for eight countries with five environmental accords through ten years<sup>9</sup>. They conclude, in harmony with the managerial school, that administrative capacity is a crucial variable when seeking to explain compliance. Supporting such findings, Simmons (1998:83) writes: “[I]acking such administrative or technical capacities, rule-consistent behaviour may simply not be within a signatory's choice set.” Taking such assumptions as their starting point, Haas, Keohane et al. (1993) argue that the main function of international agreements, next to introducing obligations to signatory states, is to facilitate a minimum performance for states who lack the resources or endowments necessary for compliance. The above discussion allows three hypotheses to be formulated:

---

<sup>9</sup> The eight countries under scrutiny were; Brazil, Cameroon, China, Hungary, India, Japan, the Soviet Union/Russian Federation, the United States as well as the European Union. The five treaties chosen were; The World heritage Convention, CITES, the International Tropical Timber Agreement, the London Ocean Dumping Convention and the Montreal Protocol on Substances that Deplete the Ozone Layer.

*H5: Compliance with the Montreal covenant is more likely when the degree of government autonomy is strong.*

*H6: Compliance with the Montreal Protocol is more likely when the degree of government capacity is strong.*

*H7: The probability of compliance with the Montreal accord increases adversely with the range of corruption in society.*

### **3.4.2 Political regime**

The importance of political regime for economic growth and social welfare is a well studied phenomenon among political scientists. The regime's role for compliance with international regimes, however, is less theorised. Beth A. Simmons (1998) introduces us to an approach she names democratic legalism. The assumption is that democratic regimes are more likely to comply with international legal obligations than autocracies. The mindset behind this argument is based on the tendency among democracies to be bound into a 'zone of law' in their meeting with the legal frameworks on the arena of international cooperation (Simmons 1998:83). Their respect for judicial processes and constitutional limitations on governing power makes democratic rulers more prone to accept rule-based constraints on their behaviour in the international sphere, especially so if the judiciary is independent. Thus, Simmons (1998:84) conclude that: "liberal democracies are more likely than are other regime types to revere law, promote compromise, and respect processes of adjudication."

Young (1979) and Neumayer (2002) also argue that democracies should exhibit stronger IEA commitment than their autocratic adversaries. Their argumentation, however, hinges on another kind of reasoning. According to Neumayer;

“(...) in democracies citizens are better informed about environmental problems (freedom of press) and can better express their environmental concerns and demands (freedom of speech), which will facilitate an organization of environmental interests (freedom of association), which will in turn put pressure on policy entrepreneurs operating in a competitive political system to respond positively to these demands (freedom of vote), both domestically as well as via international cooperation” (Neumayer 2002:140)

For Neumayer, the core of the positive link between democracy and international environmental commitment is that environmentalists can influence policy-makers. Simmons, however, sees the constraints on power by law as the decisive driving force creating a positive link between democracy and environmental compliance.

Some theorists also find it likely that democracies might not be as positive for the environment as one might think. As liberal democracies are built around a capitalist system,

private property rights and individual liberty are cornerstones in a free society. Maximising profit for one self through economic advancement is thus a self-evident right, whereas the environment might suffer severely from such a mentality<sup>10</sup> (Neumayer 2002; Desai 1998). Political economists also raise this concern, arguing that the short-lived character of a democratic government forces politicians to have dual goals; both pleasing the population and winning the next election. This narrow time prospect underlying the politicians' policy recommendations break with the necessity of long-term, non-myopic planning to combat environmental problems (Lafferty and Meadowcroft 1996:268-271; Keech 1996).

Turning to empirical studies, most scholars find a positive relationship between democracy and different environmental issues. In one of the first empirical contributions to this field, Congleton (1992) finds that when controlling for other variables, democracies are more likely to sign the Vienna Convention and the Montreal Protocol. Still, the same democracies show statistically higher methane and CFC emission rates. Such a multidimensional relationship is also supported by Midlarsky (1998). Brown Weiss and Jacobson (1998:532-534) find, in their comparative study, support for the assumption that democracies are positive for compliance. They emphasize both independent courts and public influence as important features promoting member accession. Still, evidence also show that environmental action might be hampered or delayed in democracies as government must respond to the public will, which may not be in favour of environmental commitments. Strong support for the theorised positive relationship is found in Neumayer's (2002) cross-national study of environmental commitment with four IEAs<sup>11</sup>. He observes that regardless of operationalisation, democracy enhances environmental performance. This makes him conclude: "All other things being equal, a more democratic world will also be a world with stronger environmental commitment" (Neumayer 2002:158). From this discussion the following hypothesis can be deduced:

*H8: Higher levels of democracy raise the likelihood of compliance with the Montreal Protocol.*

---

<sup>10</sup> Of course, capitalism is also found in authoritarian states. The difference is that here, private property and individual rights are often restricted to a small elite in practice, whereas it in democracies is supposed to comprise all citizens.

<sup>11</sup> The Environmental agreements studied were; the Kyoto Protocol, the Montreal Protocol, the Rotterdam Convention and the Cartagena Protocol.

### 3.4.3 Economic development

The effect of economic development on environmental degradation was first put out by the World Bank Development Report in 1992 and has since then been assigned the name ‘the Environmental Kuznets curve’, with a clear reference to the theories of income inequality put forward by Kuznets (1955), showing a U-shaped inverted pattern. Briefly, the shape of the curve shows us that as poor states become richer, they first increase environmental pressures until a turning point is reached. From this point on, the relationship changes from negative to positive and continued economic advances reduce environmental damage (Panayotou 2003:45-46). Grossman and Krueger (1995:370-371) find in their influential work that despite natural variation in the turning point for different pollutants, environmental degradation tends to start decreasing as a state reaches an income rate of \$8000 per capita. They thus provide support for a curvilinear relationship between economic development and compliance.

Despite the consensus in the early 1990s with regard to this relationship, more recent empirical assessments are critical of assigning a particular shape to all pollutants and across all countries. Re-examining the World Bank and Grossman and Krueger analyses with more recent data, Harbaugh et al. (2002) find that the evidence for an inverted-U relationship is weak. Their reassessment shows that both the turning point and the very shape of the relationship between development and different pollutants are sensitive to only small changes in the data. Panayotou (2003) points out that despite the established empirical relationship found, it is not higher income per se that creates environmental improvement. Rather, it is the responsiveness to the increased demand for protection of the environment.

Jänicke (1992:53) postulates that high levels of economic development is favourable for environmental protection because richer countries see a substantial drop in the number of people working with industrial production as the service industries turn dominant. As for environmental awareness, the author argues that education and increased leisure promote change in values and offers better preconditions for environmental mobilisation. Another empirical study finds that an increased standard of living show little discernible effect on implementation and compliance with the IEAs under scrutiny. Economic collapse, on the other hand, turned out to have a profound effect, but mainly indirectly as limited resources and rapid inflation rates made the customs inspectors more inclined to allow illicit trade in endangered species in exchange for a return service, often money<sup>12</sup> (Brown Weiss and Jacobson 1998:530-531).

---

<sup>12</sup> The study found that Cameroon and Russia failed to continue being in compliance with the CITES regime in the wake of the economic chaos that both member states experienced in the mid-1980s

There are complex processes lying behind the change from a traditional through an industrial to a modern society. Hence, one should keep in mind that whichever effect is found between economic development and IEA compliance, what lies behind is immense changes in society permeating all levels of a citizen's life. This allows us to create the following hypotheses:

*H9: Increases in economic development first decrease the likelihood of compliance with the Montreal Protocol before the effect turns and continued economic advancement sees an increased chance of reaching compliance.*

*H10: The likelihood of compliance with the Montreal Protocol increases the richer the state.*

#### **3.4.4 Legal origin**

In the literature on compliance, the importance of culture and tradition is often evaluated through a country's legal system. The assumption is that the more litigious and complex the legal system is, the more likely the member state is to be a non-complier. La Porta et al. (1999) distinguish between five types of legal origin; Common law, French civil law, German civil law, Scandinavian civil law and Socialist law. In the literature on EU-compliance, the Scandinavian civil law model is considered the most effective in their treatment of disputes as well as being respectful to international law. This record makes scholars assume that compliance with EU law happens quicker here (Perkins and Neumayer 2007:25-26).

Taking a broader approach to the importance of national legal systems, Simmons (2000) tests compliance with the International Monetary Fund (IMF). She finds that French civil law as well as Socialist law is inferior to the other systems at securing property rights, significantly decreasing a member state's probability of moving into compliance with IMF law. Furthermore, the same two groups of legal systems are expected to have the lowest level of government efficiency, further reducing the chances of compliance. In socialist law this finding is based on the argument that the extreme power of the state corrupts the bureaucracy, whereas the French civil law system's inefficiency hinges on a too autonomous bureaucracy (Glaeser and Schleifer 2002; La Porta et al. 1999:231-233).

Based on the empirical assessments of the importance of legal origin for compliance, I expect to find that member states practising Socialist law or French civil law will perform worse than Common law, German civil law and Scandinavian civil law when compliance with an environmental accord is considered.

*H11: The likelihood of reaching compliance with the Montreal Protocol declines if the member state has a French civil law or a Socialist law model.*

### **3.5 The legitimacy school**

Opposing the enforcement and management schools' understanding of non-compliance as a matter of power or capacity, scholars based in the constructivist branch of IR focus on the importance of legitimacy through social learning and persuasion (Checkel 2001; Simmons 1998). They agree with the enforcement schools' assumption that compliance is strategic, implying a matter of choice. However, the mechanism through which this happens, vary substantially. The legitimacy school considers international agreements to be social constructs, and thus they argue that compliance should be analyzed with a focus on communication between agents. Consequently, scholars assume that states comply with international law because they feel a moral obligation to do so, a so called 'logic of appropriateness' (Checkel 2001:557). Checkel (2001) introduces two causal mechanisms through which social actors internalise international norms; social mobilisation and social learning. The work on social mobilisation focuses on how domestic and international interest groups may promote pressure on state policy makers, utilising international norms<sup>13</sup> (Checkel 2001:557-560). Through the mechanism of social learning, preferences are altered in a more lasting manner, for example through compliance pull from other compliant states, acceptance of the rule-setting institutions as well as persuasion and learning through deliberative argumentations where demands and threats are absent. (Börzel et al. 2007:10-12; Checkel 2001:560-564). The theories and hypotheses presented in the following section all relate to this constructivist idea of the importance of norm-driven behaviour for understanding signatory state compliance. Public opinion, Environmental Non-Governmental Organisations (ENGOS) and green and left-libertarian parties are assumed to influence political attitudes. Moreover, the length of membership, both of the Montreal Protocol and the EU, are assumed to affect the internalisation of norms. Lastly, compliance pull from neighbouring states is theorised to affect states through persuasion and social learning.

#### **3.5.1 Public opinion**

Connected to the social mobilisation theory of the legitimacy school, compliance with an international environmental accord is considered to be higher the more concerned the population is with the problem at hand (Haas 1998:28). Brown Weiss and Jacobson (1998:29-30) point out that if the public thinks of an IEA as important and that environmental regulation is needed, then the government will be more inclined to pursue policies securing

---

<sup>13</sup> This type of "social sanctioning" against defectors have a realist touch to it as norms are not internalised; they merely constrain state behaviour.

compliance in the long run. They point out that public support for environmentalist issues suddenly became salient in the US in the early 1970s and in Europe during the second half of the 1980s. The close link between the rising concern with environmental issues and the wave of IEAs introduced in this time period should not be underestimated. Still, the environmental issue constantly fights with other political issues for attention, and the trend is that when the economy fares well, the relative importance of environmentalism grows, but as economic hardship comes, economic concerns get higher salience. Supporting these theoretical considerations, Brown Weiss and Jacobson (1998) find in their qualitative study of five IEAs that the importance of public opinion coupled with media coverage turned out as very important in persuading national authorities to act according to IEA norms.

*H12: The likelihood of member compliance with the Montreal accord increases as the population acknowledges the agreement as important.*

### **3.5.2 Montreal Protocol membership length**

Understanding the nature of compliance without taking time into consideration is hard. Most scholars working empirically with this concept see compliance with regimes as improving over time as norms are accepted and internalised, increasing the legitimacy of an accord. In this, there is a strong case to be made for the constructivist theory of the importance of social learning processes. Brown Weiss and Jacobson (1998:512, 537-542) find that: “the trend over the decades was toward greater compliance.” The mechanism through which this happens is an overall increased attention to implementation and compliance, as well as a strengthening of the supervisory mechanisms and compliance pull by other members taking a leading role. Breitmeier et al. (2006:71) find the same converging pattern of compliance over time in their study of 23 IEAs. Still, they only include what they have defined as important members in each agreement, and thus they do not represent the wide range of members in the agreements. Haas’ (1998:29) study of compliance with EU directives assigns a role to maturity as well. He argues that the political will to comply increases over time as EU institutional factors gain more influence on domestic factors. Thus, implementing a time dimension into the analysis to comprehend its importance, as well as measuring the chain of causality, seems valuable.

*H13: The likelihood of compliance with the Montreal Protocol increases the longer the time since the state ratified the agreement.*

### 3.5.3 EU membership length

One of the biggest achievements of the European Union project has been to create more interdependence between member countries. Based on this regional integration, Mitchell (2003:453) theorises that EU-members are more prone to adopt ambitious agreements and change behaviour according to IEAs more easily than less integrated states. This also follows from the fact that the EU is a contracting member to many environmental accords and has even taken on a global leadership role in some of these. As the EU signs binding treaties, it affects all member states as they are expected to live up to the standards set in Brussels (Grant et al. 2000:89-119).

Also, when considering compliance with EU-directives, a reputational effect has been introduced as a possible explanation for why there exists variation with regard to compliance. Giuliani (2003:146-147) argues that new EU-entrants are better compliers than older ones. This hinges on the assumption that new members wish to show themselves as dependable and responsible collaborators and create legitimacy among the other members. Thus, they are expected to make bigger efforts to implement and comply with IEA law. Support for this hypothesis is found in Perkins and Neumayers' (2007) study of implementation of EU environmental directives where a quantitative analysis confirm that recent entrants see less legal infringements. They thus contend that well-established members feel comfortable in their seats as they have shown themselves as reliable partners over a long period of time. Hence, they might defend national economic interests over EU environmental law.

Contradicting the abovementioned chain of causality, some constructivists inspired by the social movements' literature argue that the relationship between membership length and compliance is reverse. Checkel (2001) argues that the causal mechanism behind this assumption is that through social interaction, other member states and international actors help infringing states into adherence with the norms of the organisation and thus towards compliance. However, Checkel (2001:567-568) points out that domestic institutions influence whether the change in agent interest towards compliance happens through learning, persuasion or social sanctioning. Inherent in the social learning theory is the assumption that such learning happens over time. Thus, one should expect that with time, norms and rules are more likely to be internalised by the member states, and through this a move towards compliance should be observed. This theoretical review allows for three hypotheses:

*H14: The propensity for compliance with the Montreal Protocol increases if the member state is an EU-member.*



*H15: Young members to the EU are more likely to be in compliance with the Montreal Protocol than older members.*

*H16: The longer a state has been an EU-member, the more likely it is to comply with the Montreal covenant.*

### **3.5.4 Environmental Non-Governmental Organisations**

The role of ENGOs in global environmental politics has been heavily studied, mainly through qualitative studies (e.g. Brown Weiss and Jacobson 1998; DeSombre 2000; Jänicke and Weidner 1997; Haverland 2000). Scholars theorise that national and international ENGOs create a compliance pull on authorities and policy makers (Breitmeier et al. 2006; Axelrod et al. 2005; Brown Weiss and Jacobson 1998; Young 1979; Checkel 2001; Haas et al. 1993; Andresen and Gulbrandsen 2003; Cameron et al. 1996). Both Breitmeier (2006), Young et al. (2006:104) and Brown Weiss and Jacobson (1998) find in their empirical studies of environmental accords that the presence of ENGOs is positive for compliance. The latter scholars found that because democracies provide ENGOs with freedom to operate, transnational coalitions lobbying for improving government compliance efforts are established. This tendency has been strengthened from the 1970s to the 1990s. Haas et al. (1993), in their influential study of seven IEAs, confirms and adds further qualifications to the already mentioned empirically based assumptions. They conclude:

“[i]f there is one key variable accounting for policy change, it is the degree of domestic environmentalist pressure in major industrialized democracies, not the decision-making rules of the relevant international institutions” (Haas et al. 1993:14).

Environmental groups, then, are expected to have a positive effect on signatory state’s compliance with the Montreal Protocol. This leaves us with a new hypothesis:

*H17: Compliance with the Montreal agreement is more likely among member states where active ENGOs operate with success*

### **3.5.5 Political orientation**

In a constructivist context, the success of green and left-libertarian political parties can be held as an indication on how citizens evaluate the importance of environmental issues. The bigger these parties are, the more environmentally concerned the population. Bernhagen (2008) argues that green party popularity is a proxy variable reflecting the view a member state population has on the importance of avoiding environmental deterioration through

complying with IEAs. He finds evidence for the importance of green party popularity when considering IEA implementation, but the effect is not significant for member compliance. Furthermore, Jänicke (1992; Jänicke and Weidner 1997) finds weak results for the green seats theory. In his first review of conditions for environmental policy success, Jänicke (1992:49) renders little influence to green parties. In his more recent publishing, Jänicke and Weidner (1997:306-307) are more moderate, concluding that according to 13 case studies on environmental policy making, green parties are no necessary condition for implementation even though they may have a positive effect. They argue that part of its lacking influence is due to unfavourable electoral systems that obstructs their high voting potential from materialising.

Neumayer (2003) considers the combined strength of green and left-libertarian parties and finds, in his panel analysis of air pollution in OECD countries, that “green or left-libertarian parliamentary strength is unambiguously and robustly associated with lower pollution levels for all five air pollutants (...)” (Neumayer 2003:218). He explains the robust results by arguing that even though green parties are small in most countries, their influence stems from their success in forcing governments to take environmental concerns seriously. When concerned with left-wing party strength alone, Neumayer (2003) observes that the association with lower pollution levels are less consistent and robust, although it too point in a positive direction. Jänicke (1992:49) supports these findings with regard to the left-wing parties, arguing that: “in international comparison, left-wing governments as a whole do more for the environment than those on the right wing.” Because there are varying assumptions and findings with regard to the effect of green and left-libertarian parties, two separate hypotheses are introduced in an attempt to single out their isolated effects.

*H18: Green party popularity raises the likelihood of compliance with the Montreal accord.*

*H19: Left-libertarian executives raise the likelihood of compliance with the Montreal Protocol.*

### **3.5.6 Regional compliance**

In the democratisation literature, Doorenspleet (2004) focuses on the importance of democracy in the region for the propensity of a state becoming democratic. In the international regimes literature, such an argument has not been particularly prominent. However, Simmons (2000) argues that one should consider the effect of a regional compliance pull. In her event history analysis of commitment and compliance with IMF, she

finds statistical evidence for the hypothesis that a state is more likely to move into compliance if the contiguous countries already are compliers. Hence, she concludes that: “[o]ne of the most interesting findings is that the behaviour of other countries, especially in one’s own region, has far more influence on commitment and compliance than has generally been recognized” (Simmons 2000:832). Simmons’ findings serve as an argument for a constructivist inspired argument in close relation to the social learning theories introduced earlier. Thus, a hypothesis is developed for empirical research:

*H20: The likelihood of reaching compliance with the Montreal Protocol increases if the region the member state belongs to is one of generally good levels of compliance.*

### **3.6 Summary of arguments and hypotheses**

Table 1 summarises the theorised effects of each variable on the likelihood of member state compliance with the Montreal Protocol.

**Table 1: Expected effects on member state compliance with the Montreal Protocol**

	Expected effect
<b>Enforcement factors</b>	
Country Strength	+/-
Integration in world economy	+/-
<b>Managerial factors</b>	
Government autonomy	+
Government capacity	+
Corruption	-
Democracy	+
Level of economic development	+/-
French Civil law	-
Socialist law	-
<b>Legitimacy factors</b>	
Public attitude towards Protocol	+
MP-membership length	+
EU-membership length	+/-
Active environmental NGOs	+
Green party popularity	+
Left-libertarian executive	+
Regional compliance	+

Founded on an assumption of states as rational actors, some authors argue that *powerful states* are more likely to be bad compliers based on the international autonomy they exhibit. Others contend that powerful states are good compliers because they see this as an opportunity of showing their importance in the world. Turning to *economic integration*, opposing views also exist. Whereas some postulate that highly integrated states are better compliers because they depend on a good reputation internationally, others contend that the same countries have great incentives to keep production costs at a minimum to attract companies and investors, and this does not allow for environmental commitments.

Focusing on states' ability to comply, scholars agree that the more *autonomous a government* is from society, the more effectively it can rule, thus the faster compliance. *Government capacity*, seen through endowments and human resources, is also theorised to be positive for a state's compliance record. *Corruption*, on the other hand, is assumed to reduce a signatory state's compliance merit. A *democratic state* is taken to be a better complier than its authoritarian adversary. Likewise, resources are important to be able to comply, and even though scholars disagree on the shape of the relationship, there is consensus with regard to *economic development's* positive influence on the chance of living up to the Montreal Protocol's demands. A state's *legal system* is also assumed to impact on compliance chances. Here, Socialist law and French civil law is hypothesised to fare worse than other judicial systems based mainly on low efficiency.

The last group of theories are based on the understanding that social learning and legitimacy pushes compliance. *Public support* for IEAs is thus assumed to be positive for compliance. Furthermore, states that have been *members to the Montreal Protocol* for a while are assumed to be better compliers than newcomers because they have internalised the underlying norms. *EU-membership length* is also hypothesised to be influential on compliance both in a negative and positive way. Some scholars assume that new entrants will be preoccupied with appearing as responsible and reliable partners, and thus comply, whereas others argue that mature members are more likely to be in compliance because they have had more time to internalise laws and move into adherence. The importance of a compliance pull from *ENGOS* for compliance has also been underlined by scholars. Also, political orientation is postulated to play a role in the compliance process. Both *green seats* in parliament as well as *left-libertarians* in government are assumed to push national politics in a more environmental-friendly direction. Lastly, a theory on the importance of *regional compliance* for own adherence materialise in a hypothesis that the higher compliance rates in a region, the more likely a non-compliant state is to follow.

## **4.0 Methodology: Event history analysis**

This section will connect the theoretical discussion in chapter three with the methodological framework for the analysis. The term methodology is here understood as; “the means scholars employ to increase confidence that the inferences they make about the social and political world are valid” (Hall 2003:373). First, the choice of a quantitative model will be argued for, followed by an introduction of the case selection that proved necessary. Furthermore, the logic of an event history model will be introduced, before an introduction to the specific model of choice is presented and thoroughly discussed.

### ***4.1 Choosing a quantitative model***

To be able to draw proper inferences from the data, identifying a determinate research design is crucial (King et al. 1994:9, 118-124). In my research question I ask which factors affect member state compliance with international environmental accords. This point of departure implies two things; first, member state behaviour must be analysed, and secondly, a process across time is necessary to see how party compliance changes. The cross-unit and cross-time dimensions embodied in my research question, implies the ability of causal inference to be made. Mitchell (2004:140-143) points out that longitudinal studies able to point at causal relationships are highly needed, but lacking because of theoretical and empirical obstacles. According to King, Keohane and Verba (1994:75-97), causal processes with high certainty can best be traced through quantitative approaches. Thus, I have decided to utilise a quantitative model labelled event history analysis.

Next to enabling causal inferences, the quantitative method’s main strength is its ability to draw parsimonious conclusions from big amounts of information (George and Bennett 2005:30-31; King et al. 1994). As I include all signatory states (196 states) to the Montreal Protocol and follow them over the course of time, a large-N study is best at deducing clear and meaningful inferences based on the large sample in this thesis.

When choosing a quantitative design, I sacrifice the advantages connected with the qualitative approach. As pointed out frequently, a problem with researchers applying a quantitative method, as opposed to a qualitative design, is that they do not possess in-depth knowledge of their many cases (Hall 2003; McKeown 1999; Ragin 2004; Gerring 2004). Lacking knowledge about the units of the phenomenon under scrutiny might lead to bias. At

its worst, it might lead to model misspecifications with regard to internal validity<sup>14</sup>. Sartori (1970:1040-1046) underlines the importance of conceptualising the theories so broadly that empirical testing can be made on the whole sample, but narrow enough to be able to say something substantial about the world, as is. He thus warns against moving to such a high level of abstraction that conclusions are made that do not reflect the nature of the research object. However, I try to mitigate these shortcomings through establishing a firm research design as well as taking question of validity and reliability serious.

As a reminder, I end this section by repeating what I seek to do with my thesis, and contrasting this with what I do *not* intend to do. I wish to be able to say something about how states behave with regard to binding processes in the international sphere, more in particular, with IEAs. To absorb the notion of behaviour in my analyses, I focus on understanding what makes some member states comply early while some become laggards and others do not comply at all. This being said, I am not interested in explaining how much is achieved from a given environmental regime in total progress, nor do I regard the implementation process of an IEA as important for my study. These considerations are different processes than the evolution of compliance, and consequently should also be treated separately.

Furthermore, my focus is on parties to, or members of, IEAs<sup>15</sup>. This means that I am not conducting research on non-members to the agreements under scrutiny. I therefore follow Chayes and Chayes (1993; 1995) assuming that agreements matter inherently in pushing for change, and that the compliance processes seen within an environmental agreement differs from the pattern of behaviour seen among non-members. This assumption has been criticised under the label *the endogeneity problem*, and is according to Mitchell (2003) and Downs (1996:382-287) a problem to the entire area of research. The argument is that there exists a bias in the entire branch of research because only countries that have already become member to an agreement are put under scrutiny. If one assumes that: “[a]greements are signed only by those states that are ready to limit environmental harm – and only when they are ready to do so” (Mitchell 2003:453), there is no surprise that the research come to such positive conclusions about how IEA matter *per se*. This is a valid critique of the paradigm that should be taken into concern when considering studies within this school of research. However, my motivation for this thesis is a wish to understand what makes signatory states comply with the

---

<sup>14</sup> Adcock and Collier (2001:529) name this type of validity “measurement validity”, and define it as: “(...) whether operationalisation and the scoring of cases adequately reflect the concept the researcher seeks to measure.”

<sup>15</sup> In the IR literature, members to international accords are identified as *Parties* to an agreement. Yet, in this thesis I consistently avoid using the word party when talking about the member states. This is done to avoid confusion as party is used in this thesis with the meaning political party.

regimes they are members of. Accordingly, my focus is not on the importance of agreements alone, but on what make their members comply. Thus, I acknowledge the presented critique, but I do not feel that this thesis is up against the wall as the interest here lies in understanding state behaviour.

#### **4.2 Case selection – choosing the Montreal Protocol**

When I have chosen to study one protocol from one environmental accord, it reflects a compromise between wanting to compare as many agreements as possible and developing a framework for research that is feasible within the time and scope limits of a master thesis. According to Barrett (2003:135), there were 225 IEAs in force by the start of the new millennium<sup>16</sup>, and a protocol is only a component of an IEA. As mentioned earlier, the Montreal Protocol is part of the Stratospheric Ozone Regime which was initiated by the Vienna Convention in 1985. Then the Montreal Protocol came as the first binding framework on the Parties. Following the Protocol, four legally binding amendments have been adapted, respectively in London (1990), Copenhagen (1992), Montreal (1997) and Beijing (1999) (Andresen et al. 2008:61).

Studying only one protocol from an environmental regime may be problematic as so much information regarding the IEA is lost. However, to enable concrete regime goals that can be measured in a quantitative manner, it was necessary to move down ‘the ladder of abstraction’ (Sartori 1970) and choose one protocol to reflect the phenomenon of compliance. Since studying a single case entails a lack of representativeness, and there is a selection bias present in my choice of study object, the external validity in this study is weakened (Geddes 1990). I thus lose the ability to generalise my findings to the rest of the population of IEAs. I argue that despite this qualification, which is considered to be one of the great advantages of quantitative research, my study is important. It is one of the first attempts to analyse state behaviour within the realm of international environmental accords statistically<sup>17</sup>. Furthermore, the focus from the IR-scholars has mainly been on effectiveness of environmental regimes (Underdal and Young 2004; Miles et al. 2002; Ringquist and Kostadinova 2005; Victor et al. 1998; Young 1999; Breitmeier et al. 2006). Here the research units have been agreements, and consequently the amount of quantitative research done on effectiveness is of little use for the scholars wishing to explain member state compliance.

---

<sup>16</sup> Most of the IEAs in this review were regional, though.

<sup>17</sup> A notable exception is Bernhagen (2008).

When only studying one case, the choice of the Montreal Protocol should be properly advocated for. George and Bennett (2005:83) state: “the primary criterion for case selection should be relevance to the research objective of the study (...).” I argue that the Montreal Protocol is relevant and important because it inhabits a role as a *crucial case* within the IEA literature. As the success story of effective international environmental governance, agreements ratified in recent history have built upon a similar regime framework for maximising success (Andresen et al. 2008:67-68). Despite holding such a brand, the observed variance with regard to state behaviour is intriguing. This serves as an indication that even with the most effective IEA legal framework in place, states do not move in a monotonous process towards compliance. It is clear then, that factors external to the agreement play a role in this process, and that the institutional aspects do not paint the whole picture. Basing my analysis on the success story, I argue that the findings will give empirically rich insights into the extralegal factors driving compliance. Though the findings can not be directly transferred to other IEAs, the knowledge attained from studying compliance with the Montreal Protocol might tell us something about which national factors might be present in a compliance process and thus introduce valuable policy recommendations.

Also, I find it valuable to use one of the most heavily theorised IEAs as starting point for my statistical analysis, as it is interesting to see how the study I conduct corresponds with the theory that has focused on overall treaty compliance or effectiveness.

Lastly, there is a data aspect inherent in the choice of case. Few of the IEAs have proper reporting procedures allowing for an effective measuring of compliance. Even those regimes that do have reporting procedures are paralysed by incomplete and missing data as a result of reporting neglect in member countries. For the Montreal Protocol, data records are good as annual reports are demanded from all member states. Also, the protocol is supervised by the UNEP, providing necessary resources to provide online databases where emissions data can be accessed. These data are crucial for the operationalisation of compliance.

### **4.3 Event history analysis**

The event history analysis originated in biology, by biostatisticians analyzing the occurrence of deaths, and is probably better known under the label survival analysis<sup>18</sup>. It is a quantitative

---

<sup>18</sup> Event history analysis has also been labelled failure time analysis, hazard analysis, transition analysis, reliability analysis and duration analysis (Box-Steffensmeier and Jones 2004:2). However, this thesis will consistently use the designation event history analysis.



method, involving statistical examination of longitudinal data gathered on a set of observations (Box-Steffensmeier and Jones 2004:1-5). The starting assumption of event history analysis is that the duration spent in a state is related to the likelihood of experiencing some event at a certain point in time. In political science, the use of event history modelling has traditionally been scarce. Blossfeld et al. (1989:25) argue that the absence is due to an inability among social science researchers to use methods of dynamic analysis. Hall (2003:375) agrees, maintaining that “the ontologies<sup>19</sup> of comparative politics have substantially outrun its methodologies.” Yet, the last two decades have seen a growing body of longitudinal data, rendering event history research available. To illustrate its recent visibility, Golub (2008) finds 84 cases of event history analysis published in leading political science journals<sup>20</sup> over the time period 1989-2005. Still, the numbers are not overwhelming, signifying that even today, the method is far from common within political science research.

The very nature of social phenomena makes this model attractive, as many social science theories have an explicit or implicit interest in notions of timing and change. Criminologists are concerned with crimes, convictions and recidivism. Sociologists study job changes, layoffs and retirements, and demographers focus on marriages, divorce and deaths. Similarly, political scientists give much attention to riots, revolutions and changes of government. In each of these examples, an event occurs at a specific point in time, leading to a substantial change in the units experiencing it (Allison 1984:9-10). The longitudinal aspect often embedded in research questions makes the consideration of event history analysis natural for social scientists.

Turning to my particular research question, event history analysis is considered preferable for studying compliance with the Montreal Protocol because it can model both *whether* a member state complies with the agreement, as well as the *amount of time* that elapses before an affiliate has fulfilled the demands necessary to be labelled complier to the accord. As my thesis is mainly interested in the relationship between the length of the process towards compliance and independent variables<sup>21</sup>, it is these variables’ effect on the likelihood of reaching compliance which is in focus. Thus, the amount of time member states use to

---

<sup>19</sup> Ontology refers to “the world as it actually is” (Hall 2003:374).

<sup>20</sup> The following journals were under scrutiny: *American Political Science Review*, *American Journal of Political Science*, *International Organization*, *World Politics*, *Journal of Politics*, *International Studies Quarterly*, *Journal of Conflict Resolution*, *Comparative Political Studies*, *Journal of Peace Research*, *Political Research Quarterly* and *British Journal of Political Science*.

<sup>21</sup> The designations ‘independent variable’, ‘explanatory variable’ and ‘covariate’ will be used interchangeably throughout this thesis.

move from non-compliance to compliance is in itself not relevant for the research question posed in this thesis.

I find it justifiable to assume the process towards compliance with the Montreal Protocol to be a unidirectional process. The nature of the phenomenon I am studying makes it unlikely that parties will come into compliance with the Protocols they have signed, only to disregard them and break with the agreement at a later point in time. The Montreal Protocol creates legal limitations to production as well as consumption of the products comprised in the accord. This means that the industry is forced to manufacture new products for the market. Once the technology is “out there” to replace the pollutants in the Protocol, the demand for contaminants also decreases. In the end, there is little left of the polluting products in the market, as intermediaries have replaced them. Based on the assumption of unidirectionality of compliance, a *single-spell model* is utilised in this thesis. This means that an event can only happen a single time for each unit before it exits the sample. Thus, when a member state has reduced the emissions level enough to be considered in compliance with the Montreal Protocol, it is no longer part of the analysis because its development has been tracked and it is not considered to move out of compliance again.

#### **4.3.1 Defending the choice of model**

In the process of answering the research question asked, I considered three quantitative techniques. Next to event history analysis, the *logistic regression* and the *panel analysis* were studied. Thus, a discussion of why the event history model was chosen is necessary.

First, associate compliance with the Montreal Protocol implies a process, indicating the *dependence on time* to get accurate estimates of the effects the explanatory variables have. As event history analysis is a longitudinal model, it can account for the temporal dimension. Event history models measure time to an event, and does this through including a time dimension in the variables next to defining them as either categorical or continuous. Because of its longitudinal character, event history analysis allows for inferences about the causality between the covariates and the dependent variables to be drawn. Panel analysis also includes a time dimension in their predictors, but this analysis focuses on how covariates influence dependent variable in a temporal dimension, and fails to reflect the time to event dimension that is so intriguing with regard to event history models. Logistic regression does not allow for such a time dynamic in their analysis and is thus also considered less suitable than the event history analysis for answering the research question

Secondly, event history analysis can handle the problem of *right-censoring*. Censoring occurs in a study over time when a unit does not experience the event during the observation period<sup>22</sup>. This means that on the last year of observation we do not know how much longer this unit will survive. The logistic regression fails to distinguish the censored units from the units with an event occurring at the last point of observation. Consequently, both groups of units will be treated as having experienced the event, when in fact one of them has not. The only option for these models would thus be to omit all censored cases, creating a biased sample only including units that actually experience the event (Allison 1984:9-11; Box-Steffensmeier and Jones 1997:1415-1417). Such selection bias into the data has implications that are considered particularly grave for comparative research (Geddes 1990; King et al. 1994). As the event history model is able to include the cases of right-censoring, it is considered superior to logistic regression for my study.

Thirdly, a great advantage of some event history models is that they allow for *few specifications* to be made. Panel data, on the other hand, relies on many specifications before running the analyses. As few specifications are given from the theoretical discourse on associate compliance with IEAs, I would risk reading inaccurate conclusions about the rates and timing of events in a panel analysis framework (Box-Steffensmeier and Jones 1997:1418). Thus, an event history model is preferred over a panel design also because of its dynamic approach to study a complex process.

The great flexibility in the event history approach with regard to the time aspect, the problem of censoring and the advantage of few specifications enables insights of the entire process towards party compliance that logistic regression and panel analysis would not be able to provide. I thus argue that an event history design is the best way possible to study my research question.

#### **4.3.2 Introducing fundamental concepts: survival and risk<sup>23</sup>**

As noted above, timing and change are integrated parts of an event history analysis. Box-Steffensmeier and Jones (2004:2) argue that the notion of *survival* is implicitly present in most of the theoretical issues that social scientists grapple with. Why does civil war endure

---

<sup>22</sup> Censoring will be more thoroughly discussed in section 5.1.2.

<sup>23</sup> The notions of *survival* and *risk* reflect the model's origination in biostatistics, and consequently may seem artificial and intuitively misleading when utilizing them on different social science phenomena. However, this theoretical introduction will use them as they are part of a crucial vocabulary underlying event history analysis. In the later chapters, however, I will use more universal designations like chance and likelihood when talking about risk, and persist and remain instead of survive.

(survive) or end (fail to survive)? And why do politicians get re-elected (survive), even under unfavourable conditions? The examples imply that event history analysis is based on hypotheses formed with an implicit interest in survival. Framing research questions this way can give social sciences intriguing answers to important processes. Furthermore, survival can not be properly understood without considering the concept of *risk*. As Box-Steffensmeier and Jones (2004:3) denote, “[t]he notion of risk in political science (...) implies a conditional relationship with survival.” As a phenomenon persists – or survives – what is the risk that it will subsequently end? Or, if a member state has failed to move into compliance (survives) with the Montreal Protocol until a given point in time, what is the risk that a signatory state will move into compliance (fail to survive)?

### 4.3.3 The underlying logic of event history modelling

The three elementary concepts; the survivor function, probability density function and hazard rate, as well as their covariation, are introduced to show the nature of the event history approach. All three functions describe the same probability distribution for duration T to a failure event<sup>24</sup>. Actual survival time of a unit is here reported as *t*.

The *survivor function* reports the probability that no event is seen prior to time *t*. *S(t)* can also be thought of as identifying the proportions of units surviving beyond *t*. Consequently, all parties still not in compliance with the Montreal Protocol at the time of a given observation would be considered survivors in my thesis. At the origin time, *t=0* and *S(0)=1*, indicating that all units are survivors. As time goes by, units fail to survive, giving the survivor function a decreasing shape (or flat, if all units survive) (Box-Steffensmeier and Jones 1997:1418-1421; 2004:12-14). The survival function may be presented as

$$S(t) = P(T \geq t)$$

The concept of probability distribution involves the probability that an event T will occur at time *t*<sup>25</sup>. *Lim* signifies the limit of a function which is conditional on  $\Delta t$ , denoting change in *t*, which moves towards zero as time passes. As we assume that time is absolutely continuous, the *probability density function* of a unit looks like the following:

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t + \Delta t > T \geq t)}{\Delta t}$$

---

<sup>24</sup> Failure event is equal to failure to survive. Once an event occurs, the unit is no longer a survivor and exits the analysis.

<sup>25</sup> Time *t* is here understood as a random time within the time frame of the analysis.

The *Hazard rate* is the third component, capturing the relationship between the already mentioned survival function and the probability density function:

$$h(t) = \frac{f(t)}{s(t)}$$

This signifies the ability of the hazard function to reflect both survival and death (non-compliance and compliance) in one measure. More in detail, “the hazard function is the rate at which units fail by  $t$ , given that the duration has survived until  $t$ ” (Box-Steffensmeier and Jones 2004). Thus, the hazard rate is naturally conditional. The rate can vary between zero (no risk) and infinity (indicating the propensity for failure at that instant) (Cleves et al. 2008:8). Mathematically, the hazard rate may be expressed as:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t + \Delta t > T \geq t | T \geq t)}{\Delta t}$$

The hazard rate may take many different shapes as it changes over time. For instance, the human mortality pattern has a “bathtub hazard,” where the hazard is falling for a while after birth, followed by a long, flat line, before constantly rising again as one moves towards 100 years (Cleves et al. 2008:8).

Scholars using event history models often utilize the hazard rate,  $h(t)$ , instead of  $s(t)$  and  $f(t)$  for interpretation. I follow this practice. The reason is twofold; first, it allows me to include both the risk and survival notion in the same measure, allowing more information into the interpretation. For instance, using the hazard rate allows me to ask the question: “what is the chance that a member state to the Montreal Protocol will move into compliance, taking other factors into account?” Secondly,  $h(t)$  is useful because it is able to derive the survival and duration density functions from identifying the hazard rate (and knowing how the functions relate to each other mathematically).

As I aim to explore how explanatory variables affect the risk of associate compliance with the Montreal Protocol, a transformation of the hazard rate into *hazard ratio* is needed to make it reflect covariates’ influence on the measure. Consequently, in what follows I explain what this transformation comprises and how the hazard ratio is used to interpret coefficient effects.

#### 4.3.4 Interpreting the hazard rate

Coefficient estimates give information about changes in the *hazard rate* associated with changes in the explanatory variable. In short: “positive coefficients imply shorter survival times; negative coefficients imply longer survival times” (Box-Steffensmeier and Jones 2004:59). To assess the impact a covariate has on the hazard rate; *the hazard ratio* must be calculated. This is found by exponentiating the coefficients, showing the relative change in the hazard rate for a one unit change in the covariate (Cleves et al. 2008:131-132). For instance, a coefficient of 0.5 gives us a hazard ratio of 1.65 (because  $\exp(0.5) = 1.65$ ), whereas a negative coefficient of -0.5 has a hazard ratio of 0.61 (because  $\exp(-0.5) = 0.61$ ). As visualised through the abovementioned examples, interpretation of the hazard ratio is always relative to 1. All ratios exceeding 1 implies an increase in the hazard, and thus a positive effect, while reported ratios below 1 show a negative effect on the hazard of experiencing an event. Furthermore, to find the percentage point change in the hazard for a one point increase in the specified variable, the hazard ratio is subtracted by 1 and then multiplied by 100<sup>26</sup>. Applying this logic to the same examples as above, I find that the first coefficient (0.5) leads to an increased hazard of 65 percent (because  $\exp(0.5) = 1.65$ , and  $(1.65-1)*100 = 65\%$ ), whereas the negative coefficient (-0.5) decreases the hazard rate by 39 percent (because  $\exp(-0.5) = 0.61$ , and  $(0.61-1)*100 = 39\%$ ).

An example based on the forthcoming analyses of compliance with the Montreal Protocol shows how to interpret the coefficients. As already mentioned, a state’s political regime is theorised to influence the rate at which compliance with Protocol commitments happens. The regression shows that a one-point increase in the Polity2 covariate increases the hazard by 2,2 percent, advancing the time a member state spends before it complies with the Montreal Protocol’s emission standards<sup>27</sup>. This means that the less autocratic a country, the shorter the time before compliance.

Because of its importance for the understanding of the event history models, the difference between the hazard rate and the hazard ratio bears repeating. The hazard rate tells us the likelihood that a unit experiences an event at time  $t$ , presupposed that it has survived until this point. The hazard ratio includes the covariates in their calculation, showing how much the hazard changes from a one point increase in the specific variable. Hazard ratios are automatically reported by STATA.

---

<sup>26</sup> Percentage points are used in this analysis because we are interested in detecting the absolute difference between two numbers expressed in percent. Thus, a percentage point can be above 100 without that sounding artificial, as 0 is not nothing and 100 is not everything.

<sup>27</sup> Because  $\exp(.022) = 1.022$  and  $(1.022-1)*100 = 2,2\%$ .

#### **4.4 Applying the Event History Model to the analysis of member state compliance with the Montreal Protocol**

In this section, I discuss the event history models available for the study of member state compliance with the Montreal Protocol. After evaluating my research question up against different event history models, I present the Cox proportional hazard model as the selected tool for the analysis of compliance with IEAs. Furthermore, I present model-specific issues related to the Cox model, and consider the opportunities for combating potential problems inherent to the model of choice.

##### **4.4.1 Choosing event history model**

Behind the label event history analysis lays a variety of models available for the researcher. Event history models can be divided into two main categories: discrete-time and continuous-time models. There are two main differences between the groups of models. First, the dependent variable in a discrete model estimates *whether* a unit experiences an event or not. The continuous model, on the other hand, measures the *amount of time* spent before an event occurs. Secondly, a discrete-time formulation presumes that an event can only occur at certain, often predetermined times. To illustrate, election only happens on Election Day. However, one can easily think of processes that do not change according to predictions. Revolutions and wars can break out anywhere in time. When subjects under scrutiny behave this way, a continuous-time approach is needed (Box-Steffensmeier and Jones 1997:1423-1427). As member compliance can happen at any point in time, a *continuous time model* is needed for this study<sup>28</sup>. This means that the dependent variable reflects the amount of time spent in a social state (non-compliance) and is measured as a metric of time.

Three types of continuous event history models can be applied; a nonparametric, a semiparametric and a parametric. An important divide is whether the baseline hazard is known to the researcher or not. The baseline hazard is the hazard rate for a unit when all predictor values are equal to 0 (Box-Steffensmeier and Jones 2004:48-49). If it is not, a non- or semi-parametric model should be used. To anticipate such regularity one would need strong theories verified through models with an explanatory strength bigger than what is normally found in social science research. Imposing the wrong baseline can lead to severe bias as the shape affects the estimated coefficients (Golub 2008:531-534). As I have no strong

---

<sup>28</sup> A problem with the phenomenon under scrutiny is that even though compliance with an environmental accord in theory can be traced down to the very day, dependence on the member states' national reports only allows me to measure compliance on an annual basis. However, this is not a problem as long as the sample is of a certain size, and the incidence of ties is not too serious (see discussion of tied data under 4.6).

theoretical ground for fitting a particular shape to the data, the parametric model should not be used in my thesis<sup>29</sup>.

The main difference between the non- and semi-parametric models lies in their respective handling of covariates. Whereas a semiparametric analysis assumes that covariates shape the likelihood of an event occurring, nonparametric models compare survival experience at a qualitative level across the values of the covariates (Cleves et al. 2008:3-6, 91). For social scientists, the interest often is more with the correlation between an outcome and theoretically based covariates than with the exact shape of the relationship (Box-Steffensmeier and Jones 1997:1432; Blossfeld and Rohwer 1995:212). With this in mind, Box-Steffensmeier and Jones (2004:47) argue that:

“(…) in our view, most research questions in social science should be chiefly concerned with getting the appropriate theoretical relationship ‘right,’ and less concerned with the specific form of the duration dependency, which can be sensitive to the form of the posited model.”

The abovementioned discussion speaks for utilising a semiparametric model for my analyses of compliance with the Montreal Protocol. Its strength lies in its ability to measure the effect of covariates, but not having to deal with the issue of time-dependency included when a baseline hazard is specified. The model is labelled semiparametric exactly because of the parameterisation of covariates, and the non-parameterisation on the impact of time (Cleves et al. 2008:3-5).

#### 4.4.2 The Cox proportional hazards model<sup>30</sup>

The far most commonly used semiparametric model is the *Cox proportional hazards model* (Cox 1972)<sup>31</sup>. A version of this model will also be used for the analysis in chapter six. Its popularity hinges on its elegance and computational simplicity. The hazard rate for *i*th subject is;

$$h_i(t) = h_0(t) \exp(\beta' x),$$

---

<sup>29</sup> However, dismissing the parametric model leads to a loss in efficiency, as identifying a functional form would have provided for a better model.

<sup>30</sup> The model is also called the relative risk model (Kalbfleisch and Prentice 2002). However, this thesis uses the designation Cox proportional hazard model, or just Cox model.

<sup>31</sup> The most commonly used event history models, except from the Cox model, are the Weibull and Gompertz models as well as the exponential, gamma and log-logistic models. For further reading on these models, please confer Box-Steffensmeier and Jones (2004).



where  $h_0(t)$  is the baseline hazard function and  $\beta'x$  represents the regression coefficients and the covariates. As already mentioned, the model makes no assumption with regard to the shape of the hazard over time. Because the baseline hazard rate is left unspecified, Cox regression models do not have an intercept term. This might be unconventional, but does not restrain the strength of the model (Cleves et al. 2008:131). Also, as the baseline hazard rate is unfamiliar, *maximum partial likelihood estimation* is used to measure the parameters in the Cox model. As the baseline hazard is not directly parameterised, the rate is merely assumed to have an arbitrary shape, opening the possibility for the likelihood function to be as low as zero for the non-censored units between the events. Here, the *ordered failure times*, instead of the length of the interval between events, contribute information to the construction of a likelihood function<sup>32</sup>. Because of this, the likelihood function is not a true likelihood as the actual survival times of censored and uncensored cases are not directly incorporated into the likelihood. Based on a risk-set, denoting the set of units whose state just before  $t$  has not yet ended and consequently is not censored, the product of the conditional probability of an event happening at time  $t$  makes up the likelihood function of a unit failing (compliance occurring) (Blossfeld et al. 1989:73-74; Box-Steffensmeier and Jones 2004:51-56).

#### 4.4.3 Modelling group effects: The Stratified Cox model

This analysis is based on an extension of the Cox model, namely the *stratified proportional hazards model*. There are three different reasons identified in the literature for choosing a stratified Cox regression. First, a covariate might be fixed by the design of the research, second, a variable might have been identified as a stratification variable in previous studies, and third, a covariate that does not fulfil the proportionality requirement can be stratified to avoid related problems in the analysis (Hosmer et al. 2008:208-209; Yamaguchi 1991:108-109). Moreover, it is important that it is not considered crucial to the analysis to be able to measure the effect of the stratifying covariate on the dependent variable. For this analysis, it is mainly the research design which lies behind the decision to utilize a stratified model<sup>33</sup>. As mentioned in chapter two, the parties to the Montreal Protocol could ask for a ten year delay to reach compliance. Most developing countries asked for this, being identified as Article 5-

---

<sup>32</sup> In the parameterised models, knowing the shape of the distribution of time allows for a maximum likelihood estimation where the interval between survival times is used to construct the likelihood function (Box-Steffensmeier and Jones 2004:21-46).

<sup>33</sup> Even though the decision of using a stratified model is based on a design-specific reasoning, the variable does show signs of non-proportionality when it is included as a covariate in a normal Cox regression. This makes the argument for making it a stratifying variable only stronger.

members. With such existing heterogeneity in the data, it is likely that the hazard rate for compliance will differ by group. This, of course, is an aspect of the nature of the case under study, and hence it is not theorised. Consequently, it would be wrong to treat such a covariate the same way as the other independent variables. Still, we know that it would come out as highly significant, because it is closely related to the shape of the hazard.

To find out whether the hazard really changes against the two protocol groups, an estimated hazard function based on the range of observed failure times were created (See Figure 2 in Appendix). The graphs show that the hazards differ significantly for the two values of the Article 5 variable. This is a strong argument for allowing the hazard rate to vary by group<sup>34</sup>. To include this important aspect in the analysis, the regression is stratified by a dummy variable identifying whether the member state was given a delay or not. I thus consider the model to be improved when a stratified Cox regression is utilised as the results reflects the nature of the protocol more effectively. Also, I avoid bias in the measurements with regard to the effectiveness of the covariates, as an effect is reported taking into account the variation across the two groups of accord members (Hosmer and Lemeshow 1999:244). A residual based test confirms the improved model fit to data for a stratified model compared to a normal Cox regression<sup>35</sup>.

Statistically, a stratified analysis allows the baseline hazard to vary between the groups making up the stratifying variable. However, the coefficient effects are considered the same across strata (Cleves et al. 2008:152). Consequently, the normal Cox regression equation is relaxed in favour of;

$$h_i(t) = h_{01}(t) \exp(\beta' x) \text{ if } i \text{ is in group 1}$$

$$h_i(t) = h_{02}(t) \exp(\beta' x) \text{ if } i \text{ is in group 2}$$

Choosing the best model to reflect the issue at hand comes with a price. When estimating separate hazards for the two member groups, the partial likelihood function is formulated for each stratum. Hence, I loose the ability to compare relative duration times across groups, leading to a loss of efficiency in the parameter estimation (Yamaguchi 1991:109). Comparing

---

<sup>34</sup> Next to the stratified model, a random effects Cox regression, also called shared frailty Cox model, was also considered. The random effects model assumes that the different hazard rates are multiplicative functions of each other and that the effect of the Art.5-variable is random and unobservable (Cleves et al. 2008:156-164; Box-Steffensmeier and Jones 2004:142-148). However, since we have a variable available that we know leads to unobserved heterogeneity, it makes more sense to utilise the stratified approach. Also, when comparing model fit after having run both types of analysis on the same data, the stratified Cox model performs better than the random effects model. Consequently, the best model has been chosen, both with regard to research design and model fit statistics.

<sup>35</sup> See Figure 3 in Appendix.

the stratified regression with a standard Cox regression, it is visible that when accounting for grouped effects, several explanatory variables that were significant in the standard Cox regression lose their robustness. Thus, it seems that the loss in efficiency leads to a loss of predictive power in the stratified regression. Nonetheless, I argue that a theoretically and empirically sound model is more important than maximizing model fit and significant results.

To sum up, I end up choosing a stratified Cox model because I find it to be the model best reflecting the nature of the subject under scrutiny. As the Article 5-variable is not of theoretical interest, I do not lose any theoretically important information stratifying it. When stratifying an analysis by a dummy variable, the effects only come out as significant if the pattern of variation is found *across both member groups* to the accord.

## **4.5 Method-specific issues: proportionality and time-varying covariates**

### **4.5.1 Testing the proportionality assumption**

Even though the Cox model does not specify a distribution of the hazard rate, the covariates are expected to have the same effect over time. Thus, the ratio between  $h_0$  and  $h_i$  should be constant across time (Box-Steffensmeier and Jones 2004:48-49; Cleves et al. 2008:129-145). If the *proportional hazard assumption* is not followed, the researcher risks “enormous bias and render the estimates meaningless” (Golub 2008:537). The challenge with non-proportionality is even greater for my study as I include many time-varying covariates, and these violate the proportional hazard assumption *by definition* (Box-Steffensmeier and Jones 1997:1433; Golub 2008:538). Because of the problems connected with non-proportionality, Box-Steffensmeier and Zorn (2001) argue that testing the models for proportionality should be standard procedure.

The most frequently used test to detect non-proportional trends in the data is a residual based test called *the Schoenfeld test*, which was developed by Grambsch and Therneau (1994)<sup>36</sup>. The intuition behind the test is that if the covariates under scrutiny have non-proportional traits, the residuals will vary significantly from zero with time. Following Box-Steffensmeier and Zorn (2001:976):

---

<sup>36</sup> The Schoenfeld test can only be performed on a semiparametric Cox model in the STATA software package, giving the Cox regression yet another advantage over its parametric equivalents (Golub 2008).

“[I]f hazards are proportional (...), there should be no relationship between an observation’s residual for that covariate and the length of its survival time. Conversely, if proportional hazards do not hold, the fitted model will underestimate the hazard during those periods where the hazards are diverging and overestimate it when they are converging.”

Tests based on Schoenfeld residuals can be applied both for the model as a whole and individually to each variable. There exists no consensus with regard to which of the tests is the most powerful in detecting non-proportionality. Golub (2008:536-539) warns that the global test is not very strong, wrongly reporting proportionality even though covariate tests reveal significant violations<sup>37</sup>. Box-Steffensmeier et.al (2003:36, 45), on the other hand, claim that both tests are valuable and should be performed. To be sure, both global and covariate tests are run on all my models<sup>38</sup>.

The residual-based tests makes clear that the global test in all five models are insignificant, supporting the null hypothesis of proportionality of the covariates collectively. However, a closer look at the individually based tests reveals that *Green Party popularity* reports rather high chi square values which are statistically significant in two of five models, meaning the null hypothesis of no violation of the proportional hazard must be abandoned. In the remaining model, the same variable just misses the ten percent significance line, but as a significance level are arbitrary, I also treat the variable as non-proportional in this model.

The well-recognised solution to the problem of non-proportionality is to interact the variable with some function of analysis time (starting the year a country ratifies the Montreal Protocol) and include this interaction term alongside the original variable in the relevant models (Box-Steffensmeier et al. 2003). This is done, and a Wald chi square test is run to verify whether the inclusion of the interacted and the original variables improved model fit significantly<sup>39</sup>. The interaction term was found to be a significant improvement in all three models, and consequently will be included in the multivariate analysis.

#### **4.5.2 Capturing the functional form of time-varying covariates**

Next to securing proportionality, determining the functional form of the covariates is important to ensure a correctly specified model. This is especially important in this study

---

<sup>37</sup> His claim is based on a review of 84 articles using event history analysis from 1989 until 2005. From this sample, he replicated forty-five studies, finding grave flaws with regard to the proportional hazard assumption; especially bad was the global test.

<sup>38</sup> See Table 12 for the results.

<sup>39</sup> The results are reported in Table 6, p. 82-83.

because many of the covariates are time-varying, allowing for rather complex relationships between the covariate and the dependent variable. This thesis relies on Martingale residuals and fractional polynomials when assessing the covariates' functional form.

The *Martingale residuals test* is estimated from the Cox regression model, and is defined as “the difference between the observed event indicator, given by the censoring indicator and the expected number of events, which is given by the integrated hazard” (Box-Steffensmeier and Jones 2004:122). To evaluate whether each covariate's functional form is linear, the residuals are plotted against the independent variables. If the graph deviates systematically from 0 and does not hold a flat line, this indicates nonlinearity, and another functional form should be included (Box-Steffensmeier and Jones 2004:126). Among my covariates, *corruption* and *regional compliance* are identified as non-linear for higher values. For corruption, graphic comparison showed that utilising an exponential function created a close to linear effect on the predictor. With regional compliance, a logged transformation creates a close-to flat line<sup>40</sup>.

I also utilise *fractional polynomials* to see whether this test assumes other functional forms on other variables than the Martingale residuals test did<sup>41</sup>. Hosmer (2002:429) argues that fractional polynomials are superior for detecting inadequacies in the linear predictor<sup>42</sup>. The great advantage using fractional polynomials is that not only do they detect bad specification of the variables; they also identify the function that best reflects the nature of the predictor. Fractional polynomials provide a wide range of functional forms that the time-varying variables are tested against, adjusted for the other covariates (Cleves et al. 2008:177). The best functional form is the one with the lowest deviance<sup>43</sup> of all the possible models. If this is not statistically significant, the model with the second lowest deviance is considered and if this too fails to reach statistical significance the linear model assuming the original functional form stands<sup>44</sup>. Thus, fractional polynomials check whether they find alternative functions of continuous covariates that improve model fit significantly compared to the linear model. If the function reaches statistical significance, including it will increase model fit compared to keeping the linear alternative (Cleves et al. 2008:178-179).

---

<sup>40</sup> Also, *green party popularity* shows non-linear trends. However, the nonlinearity of this covariate has been taken into account through an interaction term. See section 6.3.1

<sup>41</sup> Fractional Polynomials was first developed by Royston and Altman (1994).

<sup>42</sup> The linear predictor in a Cox model is the logarithm of the relative hazard (Cleves et al. 2008)

<sup>43</sup> Deviance is defined as minus twice the maximised log-likelihood (Sauerbrei and Royston 1999:74).

<sup>44</sup> For technical details on the underlying logic of fractional polynomials, I refer you to Royston and Altman (1994), Sauerbrei and Royston (1999) or Cleves et al. (2008:176-179).

Running this test on all time-varying variables reveal that three of them have a more complex non-linear relationships with the linear predictor, and that taking such functional forms into account significantly improves model fit compared to the initial linear shape. *MP-membership - length* and *Country strength* are both found to have *inversely squared functions* creating the best models. For MP-membership, it is significant with a p-value of 0.10 whereas the Country strength covariate has a stronger p-value of 0.029, dismissing the null hypothesis of no better model fit with alternative functional forms. Consequently, these functions will be included in the analysis alongside their original equivalents. A squared transformation implies that the coefficient is interacted with itself, letting the effect depend on its respective level (Kam and Franzese Jr. 2007:42). An inverse function computes  $1/x$ . The effect of this computation is that small numbers are made large, and large numbers are made small. Both transformations compress the right side of the distribution more than the left side (Osborne 2002). *Regional compliance* is also found to be better reflected through a non-linear function, namely a *logged function*. This non-linear trend in the covariate is in accordance with the Martingale residuals results. Logged functions are used to allow the marginal effect of X to increase at lower levels of X and decline at higher levels of X (Kam and Franzese Jr. 2007). The logged function was found to improve model fit significantly with a p-value of 0.001.

One should not be uncritical towards such data transformations. Osborne (2002) warns that such transformations should not be utilised unless there exist clear reasons for it. Thus, the functions identified should make sense theoretically, and one should also make sure that influential observations and outliers do not influence the analysis excessively (Royston and Sauerbrei 2007). As the theories presented for *country strength* and *MP-membership – length* do not assume a certain shape of the hazard, it is hard to dismiss or confirm the findings on these premises. Influential observations were detected through their dfbeta-values, and those with unusually large values relative to the other observations were excluded for each relevant variable. Then fractional polynomials were rerun. All tests confirm the findings of the first fractional polynomials test and for *country strength* and *regional compliance* the statistical significance of their functions are even stronger. This allows us to conclude that the functional forms do not come from extreme observations, meaning they represent the data well.

## 4.6 Analysis specific considerations

*Tied data* is a problem for event history models. The event is tied if several units experience an event at the same point in time (Kalbfleisch and Prentice 2002:104-107). Because the hazard function is considered to be continuous, identically censoring of units (event occurring at the same time  $t$ ) is regarded as impossible. Tied data is problematic as it renders the partial likelihood inestimable. As I have annual time units over the course of 20 years, tied data are present in my analysis<sup>45</sup>. However, due to recent advances in software, the likelihood function can be modified to account for tied data<sup>46</sup>. There are four possible ways to deal with this problem. I have chosen to use the *Efron model* for this analysis, which is the most accurate model when the number of tied failures increases with time (Box-Steffensmeier and Jones 2004:53-59). This decision is based on descriptive statistics showing that there are quite a few member states complying the last two years of the observation period.

*Robust standard errors* are reported in this analysis. This is based on the acknowledgement that the observations in my dataset are not likely to be independent of each other<sup>47</sup>. Robust standard errors, then, account for clustering of observations and thus tend to report larger standard errors than its normal adversary. In my instance this is a more accurate test of the effects of covariates. Despite this modification of the standard errors, the coefficients are left unchanged (Box-Steffensmeier and Jones 2004:114-116; Cleves et al. 2008:317).

As mentioned earlier, a maximum partial likelihood procedure is utilised to obtain estimates of the covariate parameters<sup>48</sup>. This procedure produces measures which the goodness-of-fit of the estimated models derive from. The *likelihood ratio value* is the most commonly used model fit measure, and takes the value of  $-2$  times the difference in the log-likelihood between the model with no covariates and the full model (Box-Steffensmeier and Jones 2004:44). As this analysis utilises robust standard errors, STATA reports a pseudolikelihood value instead of the common likelihood value. This measure fails to reflect the distribution of the sample accounting for the lack of independent observations, and thus it cannot be used to compare models (Sribney 2005). The multivariate *Wald test chi squared* is

---

<sup>45</sup> As a rule of thumb, ties is not a problem if no more than 5% of the observations fail at the same point in time (Prentice and Farewell in: Box-Steffensmeier and Jones 1997:1434).

<sup>46</sup> This is another advantage of the Cox proportional hazard model. Standard parametric models have no way to handle tied data in their datasets (Box-Steffensmeier and Jones 2004:53; Golub 2008:539-540).

<sup>47</sup> Longitudinal data usually does not fulfil this independence requirement as the value on a variable one year is not likely to be completely independent from the value on the same variable the year before (Box-Steffensmeier and Jones 2004:114).

<sup>48</sup> "The function is derived by taking the product of the conditional probability of a failure at time  $t$ , given the number of cases that are at risk of failing at time  $t$ " (Box-Steffensmeier and Jones 2004:51).

another measure on model fit, which will be used to compare the different models in this thesis. It assumes an asymptotically normal distribution between the coefficient estimate and the mean vector and the covariance matrix (Hosmer and Lemeshow 1999:104), while adjusting for the number of variables added. The higher the score, the better the model is at explaining actual observations.

Several interaction terms are included in the upcoming analysis, thus hardening interpretation as well as significance testing for these predictors. *Interaction terms* are based on the assumption that the effect of one variable differs across values of another variable. Putting it differently, the effect of X on Y is contingent on the value of Z (Kam and Franzese Jr. 2007:22-23). The first type of interaction included in the analysis is *green party popularity* that is included in the analysis alongside the interacted variable *green party popularity \*analysis time*<sup>49</sup>. When interpreting an interaction term in this thesis, its effect is reported for a one unit increase in green party popularity at different values of analysis time<sup>50</sup>. Also the significance test needs to be differently carried out for interacted variables. Just like with the interpretation, the significance test depends on the joint performance of green party popularity and green party popularity\*analysis time. Thus, a univariate Wald chi squared test will be run on the original variable and the interaction variable to assess whether the interaction term is statistically significant.

The second type of interaction included in this analysis is the squared term. As mentioned, *country strength* and *MP-membership – length* are given squared functions based on their data composition. This is a special type of interaction terms where X is interacted with itself so that its effect is contingent on its own level (Kam and Franzese Jr. 2007:42). The exact same logic of interpretation and significance testing as the one mentioned above applies to this type of function.

---

<sup>49</sup> When interacting the non-proportional variable with a function of analysis time one allows the covariate's effect on the hazard of compliance to vary monotonically with length of membership.

<sup>50</sup> The effect at time t =  $\exp[\beta_1 + \beta_2(t)]$ , where  $\beta_1$  is the coefficient for *Greenseats* and  $\beta_2$  is the coefficient for *Greenseats\*analysis time* (Hosmer et al. 2008:116-120).



## **5.0 Data and variables**

When conducting research, the choice of sample as well as the operationalisation of the explanatory variables is crucial for the framework of the analysis. This chapter introduces the sample that the analysis rests on and the implications of this choice. Furthermore, the operationalisations of the hypotheses into testable variables are introduced and discussed.

### **5.1 Data**

The following section will introduce the dataset that the analysis bases itself on, succeeded by a thorough treatment of the model-specific issues of censoring and truncation in event history data. Moreover, a discussion related to the sample selection is introduced.

#### **5.1.1 The dataset**

As reflected in Table 2, I base my analysis on one dataset including all members to the Montreal Protocol. As already mentioned, many developing countries pushed for a delay for their phase-outs of CFC and halon gases. Even though it is the time until phase-out that is of interest here, the fact that they were given this delay has led to many of these countries entering the sample at a later period in time. Intuitively, it might seem that the obvious way to deal with these differences is to divide the members into two different samples. When this is not done, it is because the event history model can deal with units entering late, allowing these two groups to be included in the same sample<sup>51</sup>. Moreover, even though one could think of different processes explaining compliance in the Article-5 and the non-Article 5 groups, there is no theoretical discussion expecting such variance. I thus argue that there is no theoretical problem including all members in the same dataset, maximising the number of observations.

---

<sup>51</sup> See the discussion of left-truncation in the following section.

**Table 2: Overview of the dataset**

Analysis of party compliance with...	Time period	Number of compliance incidents	Average number of member years until compliance	Number of countries
...the Montreal Protocol	1989 - 2008	118	9,5	154 <sup>52</sup>

Before going into details on data-specific issues, a clarification is needed. By *member state*, I refer only to states that have *ratified the treaty*<sup>53</sup>. This means that states that have signed the Montreal Protocol but not ratified the agreements are excluded from the analysis. In IEA bargaining processes signature implies nothing more than an intent to seek ratification (Barrett 2003:147). It is first when the protocol is ratified and has entered into force<sup>54</sup> that a state is legally bound to act in accordance with the treaty text. For the Montreal Protocol this happened on January 1<sup>st</sup>, 1989.

### 5.1.2 Model-specific issues: censoring and truncation

An issue specific to the event history analysis is *censoring*. This is considered to occur: “whenever an observation’s full event history is unobserved” (Box-Steffensmeier and Jones 2004:16). *Right-censoring* is when a unit does not experience an event within the observation period. This type of analysis is the perfect tool for taking care of such instances, as the model is able to distinguish between the units having experienced an event and the ones that have not (Cleves et al. 2008:30-31). The contrasting case of *left censoring* is experienced when; “the event occurred at some time when the subject was not under observation” (Cleves et al. 2008:43). In this thesis, left-censoring refers to units where compliance has occurred before

<sup>52</sup> As of 25 November 2009, there were 196 Parties to the Protocol (The Ozone secretariat 2009). However, Andorra, Iraq, San Marino and Timor Leste are excluded from the sample as they became members on the last point of observation or later. Thus, we will not gain sufficient information to draw meaningful inferences from these units. Furthermore, 32 countries were omitted from the sample because they have less than 500, 000 inhabitants and therefore are not included in the Polity Project’s democracy scores (CIA World Factbook 2010a). See Table 10 in Appendix. Also, Ireland and Portugal are omitted because the Ozone Secretariat has never registered any emissions information on them. Consequently, I am not able to measure their compliance record. EU is a member to the Protocol as well, but is naturally excluded. Lastly, Belgium, Denmark and Solomon Island are excluded from the sample because they are left censored cases (see 5.1.2).

<sup>53</sup> Ratification normally requires the approval of a state’s parliament and is considered a formal commitment from the state under question to implement IEA law (Barrett 2003:147).

<sup>54</sup> A protocol normally enters into force 90 days after the ratification of a certain amount of signatory states. Since the protocol considered here was opened for ratification quite early in 1988, some states ratified the Protocol already this year. When the starting year for my analysis is 1989, this is because I consider the day of entry into force as the first year of analyses. I do this because the Protocol text is not legally binding on the members prior to the entry into force, and therefore their ratification may be considered nothing more than a vote of consent before it actually becomes binding.

the period of observation starts<sup>55</sup> (Yamaguchi 1991:4-8). This is an unresolved problem for event history models, leaving me with two options. Either, I can treat the left-censored cases as having experienced the event at the first year of observation in the study period, or the left-censored units must be omitted from the analysis. The former option risks biased coefficients if there exists discrepancy between the covariates' values in the attributed years and the years of actual change. The latter option means a loss of information by excluding units from the sample. However, sample bias is avoided (Allison 1984:56-57; Mooney and Lee 2000:231). Because discarding the left-censored units is superior at avoiding bias, it has become the standard solution to this problem. I follow this advice, excluding left-censored units from the analysis<sup>56</sup>. Thus, the sample represents *the members to the agreement that were not in compliance with the Montreal legal framework before ratifying the accord*.

*Truncation* is another issue characteristic to the event history analysis. Box-Steffensmeier and Jones (2004:16) consider *left-truncation* to emerge in event history datasets “when history prior to the first observation point is unobserved.” More in detail, the primary concern is that some units might enter the risk set before the initial observation point  $t_0$ . Consequently, important information about their event history is lost. Left-truncation is absent in my data as units are only considered to enter the risk set the year they ratify the agreements under scrutiny. Before this, they are simply not considered to be in danger of experiencing compliance because they have not legally committed themselves to comply. *Right-truncation* is evident in units that have entered the risk set but are not observed throughout the observation period. The model treats such cases the same way as right censored cases, assuming that failures are certain to occur sometime outside the observation period as time extends to infinity (Cleves et al. 2008:36). Austria, Finland, Greece and Spain are right-truncated in the Montreal Protocol. The great advantage with event history analysis is that it is able to treat truncation in the data<sup>57</sup>.

### 5.1.3 Sample selection

The ideal type in quantitative studies is a sample including the entire population, meaning all states at all times. As this is usually impossible, based both on lack of data as well as lack of

---

<sup>55</sup> Box-Steffensmeier and Jones (1997:1422) define left censoring slightly different, considering left censored units as incidents where the point of beginning is not known, whereas the point of failure is. This corresponds with what I define as left truncation.

<sup>56</sup> The omitted left-censored units are Belgium, Denmark and the Solomon Islands.

<sup>57</sup> In contrast, the standard OLS regression model would treat left-truncated observations as having equivalent entry times to all non-left or non-right-truncated observations.

resources, random samples where all empirical units have the same probability of being included have become the feasible ideal. Even though most statistics assume random samples, this is not always easy to fulfil (Grønmo 2004:92-98). Since the universe of members to international environmental accords is so big and complex, a random sample is hard to obtain. To reduce complexity and the costs of data collection, the data are based on a *clustered sample*<sup>58</sup>, where one cluster is first defined from the entire universe of cases (all IEAs), and then units within this cluster are identified as research units (all parties to the Protocol)<sup>59</sup>. (Pennings et al. 2006:61-62). A consequence of a clustered selection is that this study cannot generalise<sup>60</sup> its findings to the broader population of IEAs<sup>61</sup> (Grønmo 2004:97-98). First, all non-random samples entail a loss in security and precision as the random variance between the universe and the sample increases. Secondly, since I only consider one cluster, the generalisation potential is small as the representativeness of the sample decreases. However, this study is exploratory and consequently serves an important function providing comprehensive research in a format previously understudied within this paradigm.

While the case selection in chapter four decides the maximum amount of data possible, access to data also influences the final sample which the analysis rests upon. As already mentioned, not all members to the agreement are included in the analysis, some of them because of lacking data, other because of data-specific concerns. Nonetheless, the evictions are not considered to lead to bias in the analysis as states on all continents are affected.

## ***5.2 The dependent variable: Time until compliance with the MP***

This thesis aims at understanding the determinants of an associate's compliance with the Montreal Protocol. More precisely, within the framework of an event history analysis, the dependent variable measures how much time (in years) elapses from a member state enters the sample (ratifies an agreement), and thus becomes at risk of compliance, until the agreement is complied with by the same state, if it complies at all. Quite a few of the countries included in this analysis enter the analysis at the start of the observation period, here understood as the first year of ratification. For the Montreal Protocol the first year of

---

<sup>58</sup> It has also been labelled multi-level sample (Pennings et al. 2006:62).

<sup>59</sup> Even though I analyse all units within the Montreal Protocol, my study is not a population study. The universe of units identified in my research question is 'all parties to IEAs'. A population study based on my thesis would therefore include all parties to all IEAs, and not only the parties to one protocol.

<sup>60</sup> The generalisation potential is also commonly referred to as the 'external validity' of the data (Pennings et al. 2006:67).

<sup>61</sup> Consequently, significance tests will only be interpreted as indicating the likelihood that the results are not due to random measurement errors of coincidence.

observation is 1989. However, many of the Article 5 states to the Montreal Protocol enter the sample later in the observation period. As already mentioned, these incidents of left-truncation create no problem for event history modelling.

To understand what is implied in the event, defining and further operationalising *compliance with the Montreal Protocol* is necessary. In the absence of previous quantitative research on compliance<sup>62</sup> (to my knowledge), the operationalising and coding of this variable has been done by the author. I consider an associate to be in compliance with the agreement if it successfully makes the changes necessary to fulfil the goals it has obliged itself to within the protocol framework. Considering the Montreal Protocol, the non-Article 5 members have committed themselves to *phase out (100% reduction) CFC and halon gases* by January 1<sup>st</sup> 1996 and January 1<sup>st</sup> 1994 respectively. The Article 5 states negotiated a delay, and were supposed to phase out CFC and Halon gases by January 1<sup>st</sup>, 2010. The baseline year on which reductions were to be calculated was set to 1986.

The Montreal accord secretariat considers a member state to be in compliance when it reduces the amount of pollutants in question within the deadline set in the agreement. In this thesis, however, *compliance is coded based on the reduction criteria only*. This means that a party will be coded to be in compliance with the given Protocol regardless of whether it complies within the deadline or not, as long as the event happens in the observation period. Such an operationalisation of compliance reflects this thesis' interest in identifying real behavioural change. Problematic as it may be, it is not possible to include the deadline information in the measure of the dependent variable<sup>63</sup>. Yet, I argue that this can be defended, as IEAs initiate complex processes of behavioural change where the success should not be measured solely by whether they comply on time, but also how the change in behaviour materialise in visible environmental quality data.

Based on the operationalisation above, the compliance variable is coded as a dummy variable, with the value 0 for non-compliance and 1 for compliance. I utilise data from the UNEP Ozone Secretariat web page (United Nations Environmental Programme 2010). The coding is based on the variable '*ODS Consumption in ODP Tonnes*<sup>64</sup>', for CFC and halon

---

<sup>62</sup> Honorable exceptions are Simmons (2000) and Bernhagen (2008).

<sup>63</sup> A better compliance measure would probably be an evaluation of how international law materialises to national policies, and to evaluate whether these are sufficient to create the necessary change in behaviour, thus detecting members that comply without actually making an effort. However, such a measure would entail an enormous job collecting information for each member state. Thus, a proxy of actual emission reductions is used, assuming that national policies are behind the reduction records.

<sup>64</sup> ODP is short for Ozone Depleting Potential, and refers to the amount of ozone depletion caused by a substance. This weighting of raw emissions numbers is done to reflect the impact on ozone of a chemical substance compared to the impact of the pollutant CFC-11 (the basis), which is considered to be equal to 1. CFCs have

gases, and covers the period from 1989-2008. The Ozone Secretariat (2004) operationalises 'consumption' as "production plus imports minus exports of controlled substances minus destroyed quantities minus feedstock uses of controlled substances." In this measure both consumption and production are accounted for in the same variable. To take into account the different starting points for the member states, the relative reduction in ODS production and consumption is considered. This means that the dependent variable is measured based on the percentage *reduction relative to the production and consumption of ODS* in the baseline year. This way, a behavioural dimension is included in the measure as it is not insensitive to the various starting points<sup>65</sup>.

Even though 100 percent reduction is the requirement in the Montreal Protocol, the parties have been coded to experience compliance the year they have reduced 95 percent or more of the registered consumption in the baseline year. I do this because the Protocol allows for exemptions for "laboratory and analytical uses" and "essential and critical use". A brief calculation of exemptions given to several countries in 2005 and 2006 showed that they never exceeded 5 percent of the baseline consumption. Consequently, I consider this to be a proper way of taking this qualification into account in the measure.

When basing the coding on national reports written by the parties themselves, the potential problem of lack of reliance and thus low internal validity<sup>66</sup> should be discussed. Reasons for potentially inadequate estimates may be found in different methods utilised among parties for collecting emissions data, errors of emission, over-reporting, under-reporting, failure to report as well as mis-categorisation of substances<sup>67</sup>. To evaluate the data, the United Nations Statistics Division (2010) has run a test on the discrepancy between global and national figures of emission levels on the Montreal Protocol from 1997 to 2005<sup>68</sup>. They find the average error level to be 6% annually for that time period. This is an acceptable error

---

ODPs ranging from 0,6 to 1, whereas halons have ODP values up to 10 (United Nations Statistics Division 2010). The numbers I operate with to code the dependent variable therefore are the natural emissions data reported from the parties, weighted with the ozone depleting potential of the substances reported.

<sup>65</sup> Inherent in every IEA is the fact that it always takes less for some countries to comply than others simply because they are not as severely struck by the problem at hand. This mean that the best compliers can be the member states where the problem is of minor importance, whereas the laggards may be the countries actually making a proper effort to combat the environmental issue of the IEA. This means that for some states minor changes in environmental behaviour is necessary, whereas it for others requires a great deal of actual changes.

<sup>66</sup> Internal validity is concerned with how well the measures represent the theoretical concepts are supposed to reflect (Midtbø 2007:25).

<sup>67</sup> Deliberate refusal or outright manipulation of data are rarely the reason behind reporting failures, an ILO working group has concluded (Brown Weiss and Jacobson 1998:46).

<sup>68</sup> The study is based on the assumption that globally, net exports should equal net imports, given that all national data reported are accurate.

level, indicating a high degree of validity. To sum up, relying on self-reporting by the parties is potentially a problem, but here the internal validity of the measure is found to be good.

As mentioned earlier, the previous research on environmental regimes has been mainly qualitative. Several authors argue that the main reason for this methodological imbalance reflects the challenges in measuring a non-observable process like compliance (Underdal and Young 2004). Brown Weiss and Jacobson (1998:4-5) treat this more in detail, contending that even though member state compliance is measurable in principle, precise measurements are elusive.

“(…) [O]ne can tell if the production of a prohibited substance such as chlorofluorocarbons has been discontinued. But measuring emissions resulting from changing refrigerators in automobile air conditioners is much more problematic. And how should an occasional emission be weighted in relation to the overall cessation of production?”

Well aware of the challenges it poses to study quantitatively something that cannot be properly observed, I have taken my precautions. The operationalisation of compliance shows that I combine an environmental indicator with a behavioural indicator. I consider net reduction of ODS, but it is measured relative to member states' starting point in the baseline year. Doing this, I follow Mitchell (2004) and Haas et al. (1993) who argue that environmental improvement and behavioural change are appropriate metrics to use when seeking to explain variation in compliance. Moreover, Mitchell (2004:125-126) argues that a focus on change in behaviour over time (in my case seen through reduction in ODS) is the preferred way to understand compliance. As international regimes initiate complex processes of change, the results in the environmental indicator might be visible long before as well as long after deadline. Thus, by analysing net reductions over time, occasional emissions are likely to be disclosed during the coding of the variable. Consequently, only long-term efforts to reach compliance have been identified as compliance according to the coding procedures.

### ***5.3 Operationalisation of the independent variables***

This section will present the operationalisation of the factors that were hypothesised to influence compliance in the theory chapter. In general, the reliability of the data is considered to be good, as I utilise a quantitative method and most measures are derived from well reputed databases, with codebooks allowing for accuracy and consistency, ensuring replication possibilities. The question of internal validity will be dealt with consecutively if it is considered important to discuss for the variable at hand.

### 5.3.1 Enforcement factors

A state's *population status* as reported by the World Databank (2010) is used to measure *country strength*. To ease interpretation, the actual population numbers are reported per million inhabitants. Even though it is questionable whether a state's population is a sufficient measure on strength, it is commonly used in the literature. Thus, I also utilise this, despite its reduced validity. As a fractional polynomials test showed that this variable had a more complex relationship with the log hazard, it is presented through an inversely squared function. This is done to ease interpretation so that hazard ratio and standard errors make intuitive sense. Influenced by Perkins and Neumayer (2007:28), the natural logarithm has been used as it is not expected that the size of a population increases linearly with its non-compliant behaviour.

*Integration in the world economy*, or trade dependency, is measured by the size of a country's international trade. Data are gathered from the World Databank (2010), and the variable *Merchandise trade as percentage of GDP*. The trade concept is referred to as the sum of imports and exports of goods and services. To ease interpretation and not make the distance between values so small that its effect seems diminutive, the measure has been divided by 10.

### 5.3.2 Managerial factors

*Government autonomy* is measured through the variable *number of veto players*, derived from the World Bank Database of Political Institutions (Keefer 2009). The indicator relies on a check and balance system when coding the units. Based on an index of electoral competitiveness, those that score less than 4 are considered not to have electoral competitiveness. These are automatically coded to have 1 veto player, as only the chief executive wields a check when the legislature is not freely elected. However, those that score 5 or more are considered to be sufficiently competitive. Then, the amount of veto players is incremented by 1 if; a) there is a chief executive, b) the chief executive is competitively elected and c) if the opposition controls the legislative. Furthermore, the different regime types also affect the amount of veto players possible for a state. In presidential systems, the amount of veto players increases by one; a) for each chamber of the legislature<sup>69</sup> and b) for each party that is allied with the President's party, but has another ideological orientation. For parliamentary systems, three features affect the amount of veto players. The amount increases

---

<sup>69</sup> Unless the executive party has a majority in the lower house AND a close list system is in effect (as this implies stronger presidential control).



by one; a) for every party in a government coalition<sup>70</sup> and b) for every party in government that is closer to the opposition party than to the major government party when economic issues are discussed. Lastly, the prime minister's party is not considered a veto player if there is a closed rule in place<sup>71</sup> (Keefer 2009:14-19). Where data are missing for a unit, but there exist data before this year, the last registered value will be used to fill the missing years. This variable is evaluated to hold high validity.

In the discussion of *government capacity* in the theory chapter, emphasis is put on financial endowments and human resources as crucial for a state's ability to act. To capture this understanding in the analyses, the *Human Development Index (HDI)*, published by the United Nations Development Programme (UNDP) (2009), is utilised as the measure of government capacity. Embodied in this measure is the confession that it is the manner in which the country spends its wealth, and not the wealth in itself, that creates development<sup>72</sup> (ul Haq 2003). This measure has three components; health, knowledge and standard of living. Health is measured as life expectancy, based on population health and longevity. Knowledge is operationalised as a weighted education indicator where the literacy rates accounts for 2/3 of the variable's value and the total gross enrolment of primary, secondary and tertiary schools account for the last third. Lastly, standard of living is based on the natural logarithm of GDP per capita Purchasing Power Parity (PPP), to reflect the reduced importance of income with enlarged GDP. Even though this indicator is originally a measure of the general development of a society, I argue that the literature on government capacity moves close to this understanding. Consequently, I contend this measure to hold good validity. As annual data was not available until 2005<sup>73</sup>, the last registered values are used to fill in missing values if there exist data previous to the missing period. HDI is considered a rather constant phenomenon, thus this procedure should not be problematic. Also, the original measure varies from 0 to 1; however the measure has been recoded into values between 0 and 100.

The variable *corruption* relies on the World Bank measure *Control of Corruption* found in the Worldwide Governance Indicators project (WGI) (The World Bank Group 2009a). The measure ranges from -2.5 to 2.5, with higher values corresponding with more control of corruption (Kaufmann et al. 2009). The index is made up by specialist assessments

---

<sup>70</sup> Given that the party is needed to maintain a majority in the parliament.

<sup>71</sup> This is because in such settings, the prime minister is presumed to have full control over the party.

<sup>72</sup> Following this line of argument, human development is assumed to depend on 1) the formation of human capabilities (improved health, knowledge and skills) and 2) the use people make of their acquired capabilities (employment, political affairs and productive activities) (ul Haq 2003:18-19).

<sup>73</sup> Previous to 2005, data were gathered every 5 years from 1980 onwards (United Nations Development Programme 2009).

and opinion surveys from different institutions that all focus on corruption as the abuse of public office for private gain<sup>74</sup>. This variable has two shortcomings. First, data is only reported from 1996 onwards, in biennial intervals until 2002, followed by annual data. This means that there is no data for the early observation period<sup>75</sup>. To avoid many missing values in the dataset on this variable, a separate model will be run only including the years from 1996 onwards when assessing the impact of corruption on compliance. Secondly, the validity of the measure is reduced as there exists a divide between what one wants to analyse (corruption) and what is actually measured (perception of corruption). Corruption is a sensitive issue, and one can easily think of numerous reasons why a respondent would not report its actual presence in society. Consequently, the use of a perception index reflects the difficulty of measuring the phenomenon directly. However, as very few incidents of corruption are reported, objective measures would also be nothing more than imperfect proxies. I therefore argue that this measure can reflect the situation in a realistic manner.

A note should be made on why Transparency International's Corruption Perceptions Index (CPI), which is the most frequently used measure of corruption, have not been applied to this study. First, the World Bank measure included 154 countries in their first round in 1996, whereas CPI only covered 41 countries in their first year, 1995. As the scope of this thesis is broad, missing values would have become a problem using CPI. Secondly, the basis for longitudinal comparison of the CPI measure is challenged by low validity. Transparency International report that their focus is not on changes over time, and has consequently edited their sample as well as their methodology over the course of time (Lambsdorff 2006:3). In contrast, the WGI-project focus on the advantages of comparisons across time, and changes to their samples have been done in a way so that the measures are still comparable (Kaufmann et al. 2009:19-22). Thirdly, the WGI-measure includes annual aggregate measures of corruption, whereas the CPI index for a year is made up by their scores in prior years. When using averages over time, the values are not independent of each other, causing unavoidable problems with autocorrelation. Lastly, the correlation between the two measures is close to

---

<sup>74</sup> These six sources make up the measure: World Bank's Country Policy and Institutional Assessments, Political Risk Service's International Country Risk Guide, Global Insight Business' Conditions and Risk Indicators, Global Insight's Global Risk Survey, Economist Intelligence Unit's Country Risk Service and Democracy Index and World Economic Forum's Global Competitiveness Survey (Kaufmann et al. 2009:39)

<sup>75</sup> An alternative corruption measure provided by the Political Risk Services Group (2010) exists and includes corruption measures for the entire time period. Nonetheless, these data are not available as the company owning it only share data in exchange for money, and the university made it clear that they would not pay for this access.

one, implying almost complete correlation<sup>76</sup>. This means that I do not lose valuable information about corruption when choosing the WGI over the CPI measure.

*Level of democracy* is measured using the *Polity2 score* (Marshall and Jaggers 2009a). The Polity IV Project provides coding of democratic and autocratic “patterns of authority” in all independent countries with a population exceeding 500 000. The project is extensive, including regime data from 1800-2008, and has become leading in collecting data on regime characteristics (Marshall and Jaggers 2009b)<sup>77</sup>. In the original Polity score, democracies and autocracies are measured separately, based on expert evaluations of their competitiveness of executive recruitment, openness of executive recruitment, constraints on the Chief Executive and competitiveness of political participation. However, the acknowledgement that regimes can hold democratic as well as autocratic characteristics led to a synchronisation of the two scores. Later, the Polity2 variable was added to render time-series analyses possible. The variable is a modification of the combined annual Polity score, ranging from -10 (strongly democratic) to +10 (strongly autocratic) (Marshall and Jaggers 2009b).

Because data on democracy have become so available and central in much social science research, some critical remarks with regard to validity prove necessary. An important critique against the democracy indexes in general is that they aim at making an unobservable latent variable observable through different indicators, but do not seem to treat the chance of measurement error inherent in the selection process. In their comparison of nine democracy indices, Munch and Verkuilen (2002) find that only two provide a detailed justification for their indicators. Moreover, Hadenius and Teorell (2004) find large discrepancies at different levels of the democracy scale when comparing five leading democracy indices. Moving to the specific issue of the Polity index, the conceptualisation has been criticised for focusing solely on contestation and not including an inclusiveness aspect as theorised by Dahl (1971). Another problem is the redundancy present in the index attributes<sup>78</sup>. This problem is never explained, and consequently many incidences of double counts are seen without any reason put to it. Moreover, its aggregation rules have been criticised. In short, the different indicators are considered to be equally important measures of democratic quality through a weighting

---

<sup>76</sup> Running a bivariate correlation analysis shows that the correlation between the two corruption indexes is ,971, the correlation being significant at a 1% level.

<sup>77</sup> The Freedom House indicator for democracy and civil rights is not utilised in this thesis due to the limited accuracy of the rating over time. As Neumayer (2002) point out, changes in the score over time might reflect global changes rather than national institutional transformations.

<sup>78</sup> The redundancy problem is evident in the Polity IV and Hadenius indices. First, they use the attributes competitiveness and regulation of participation, which reflect the competitiveness of a regime. Also, the indices competitiveness and openness of executive recruitment both measure competitiveness in connection with offices (Munck and Verkuilen 2002:14).

scheme that is not justified. The attributes are added together, assuming that this reflects the nature of democracy. Thus, a moving from 1 to 2 on indicator  $j$  is assumed to have the same impact as an increase from 1 to 2 on indicator  $k$  (Munck and Verkuilen 2002). Furthermore, Treier and Jackman (2008:205-206) point out that the assumption of independence between the indicators is broken. The pattern found in the interaction between the indicators clearly show that few of the combinations of values possible are actually observed, thus lowering randomness between attributes. Comparing different democracy measures, the authors also find that the cluster of cases at top and bottom of the Polity scale has comparatively high level of measurement errors. When these countries are obviously very different<sup>79</sup>, despite being given the same score, the precision of the Polity estimate must be treated carefully. Despite this critique, the Polity index is one of the democracy indices that best treat the problems connected with conceptualisation, measurement and aggregation. Also, Munch and Verkuilen (2002:29-30) end up finding high correlations between the democracy measures, indicating that despite different procedures, they seem to be measuring the same.

*Level of economic development* is measured by *GDP per capita PPP in constant 2005 US dollars*. Data are gathered from the World Bank (2009b). I have registered the data per 1000 US dollars to make the numbers easier to manage. Also, the natural logarithm of the GDP measure is taken as it is assumed that the impact of each additional dollar decreases as income increases. As mentioned in chapter three, the theory of the Environmental Kuznets Curve argues for a squared relationship between economic development and the success of environmental policy. Because of expectations of collinearity, pre-tests (not reported) on the linear variable and the squared term for different samples are run. I find that the squared term fails to reach joint significance across all models, whereas the linear model performs considerably better. Hence, the squared version is not included and H9 is not further tested.

Based on La Porta et al. (1999) a nominal operationalisation of different legal systems has been reported in the Quality of Government dataset (Teorell et al. 2009). The original variable has been recoded into two dummy variables to reflect the theory that Socialist law and French civil law are unfavourable when compliance with IEAs is considered (Simmons 2000). To create more flexibility in the estimations and see whether they yield interesting results, additional dummies considering Common law, German civil law and Scandinavian civil law are also included. When dummy variables are exclusive and complete like here, one

---

<sup>79</sup> In their analysis, Treier and Jackman (2008:210) find that countries as diverse as Mongolia, Costa Rica and Papua New Guinea are considered to have democracy levels that are indistinguishable from the United States, based on data from 2000.

dummy must serve as a reference category to avoid perfect collinearity (Cleves et al. 2008:172-174). Here, I make the Scandinavian civil law dummy the reference category, omitting it from the remaining estimations.

### 5.3.3 Legitimacy factors

World Value and European Value Surveys (WVS/EVS) were used to gather data on *public attitudes* towards the government with regard to environmental issues. Data from four waves made up the background for the data, asking respondents to give their opinion about the statement “*Government should reduce environmental pollution*” (European Values Study Group and World Values Survey Association 2006:125). However, there are too much missing data, and a minimum of the countries under study have data covering the entire observation period. Consequently, it is not methodologically justifiable to run this variable in the analysis. Excluding variables is problematic when there exists theoretical expectations considering their importance. Nonetheless, too many missing values create problems for the entire sample, even though sophisticated methods have been developed to handle missing data<sup>80</sup>, I find excluding the variable to be the soundest solution. Thus, I conclude that public support will not be further dealt with in this thesis.

*MP membership - length* is a numeric variable coded with information from the secretariats of the Montreal Protocol (The Ozone secretariat 2009). The year an agreement enters into force on a party is coded as the first member year. Following a fractional polynomials test, the covariate exhibited an inversely squared function with the log hazard. Thus, this variable is presented through an interaction term in the analysis.

The three hypotheses introduced in the theory chapter necessitate two variables to be reflected. First, an *EU-membership* dummy is included to reflect H13. Then, a numeric variable identifying *EU membership - length*, counted from the year the member countries enter the EU, is included to reflect H14 and H15. The coding of the variables is based on information taken from Perkins and Neumayer (2007). Because of collinearity between the two measures, as well as with the theoretical expectations being strongest for the numerical covariate, the *EU-membership* dummy is excluded from the upcoming analysis<sup>81</sup>.

---

<sup>80</sup> See Honaker and King (2010) assessment on “multiple imputation.”

<sup>81</sup> It is always unfortunate to leave out theoretically justified variables from an analysis. Nevertheless, running the dummy variable for EU-membership through the models (not-reported), the predictor fails to reach significance in all of them. Thus, we can be certain that the null-hypothesis of no relationship between EU-membership *per se* and chance of reaching compliance stands, and that it was not of importance to the analysis.

The importance of *active ENGOs* pushing for compliance is hard to measure. Ideally, financial resources and membership base should make up this measure, but such data are not easy to get hold of. Binder and Neumayer (2005:531) utilize the number of ENGOs per capita as a proxy for an optimal measure where budgets and membership base is considered. As opposed to Binder and Neumayer's limited sample, my global sample further complicates such a measure as proper information of ENGOs in all countries in the world is hard to acquire. Alternatively, measuring the number of environmental NGOs at the national level could be a good measure. However, Roberts et al. (2004:28) point out that this measure is problematic for two reasons. First, the word "environmental" has become a rather fuzzy concept that is hard to measure strictly. Secondly, NGOs in the developing world are often cross-issue oriented and do not merely treat the environment alone. Consequently, a proxy indicator, based on the Yearbook of International Organisations and reported by the UNDP Human Development Report (2002:42-45), measures *the number of active NGOs* and is used to reflect ENGO strength in civil society. This has been measured for year 2000 for all countries; consequently this variable is treated as constant in the analysis. Although not optimal, this is the best available measure on global ENGO activity. The assumption underlying this selection is that the more NGOs there are in a society, the more ENGOs one would expect to find in a nation. This measure is problematic. First, one might expect the number of NGOs to vary with the size and geographical placement in the world, creating bias in the estimates. Secondly, a high number of NGOs in society does not equal presence of ENGOs. However, we know that where few NGOs are found, chances are high that environmental organisations are also prevented from operating. As this indicator is a proxy, the validity of this measure is weakened, and consequently the results should be interpreted with care. To ease interpretation, the variable is reported per 100 NGO in a society.

The variable *green party popularity* is measured through a global measure that estimates the *distribution of seats in parliament* gained through elections by Green parties<sup>82</sup>. Information about green parties worldwide are drawn from the Global Greens Network (Global Greens 2010), whereas election results are found in the PARLINE database managed by the Inter-Parliamentary Union (2010). Green parties are first and foremost existent in the developed world, as a result of post-materialist values theorised to coincide with increased wealth (Inglehart 1997). However, it turns out that Green parties are prominent in quite a few

---

<sup>82</sup> To evaluate which parties should be included under the label 'Green parties' I primarily base my evaluation on the parties that are members of the Global Greens network (Global Greens 2010). But not all green parties are affiliates here, and to avoid bias all parties that include the words 'environment', 'ecology' or 'green' in their party name are considered a green party.

developing countries as well, making their inclusion in this variable important. Where information about green parties is unavailable (mostly authoritarian regimes), additional sources are used. In most of these states, political parties turn out to be banned. Consequently, these were given the value 0, the same value as for states that have multiparty government, but where the greens are not represented. This is a problem theoretically, as green parties have the opportunity to enter the parliament where parties are allowed, whereas the situation is completely different in states where parties are banned. Nonetheless, I argue that this is an acceptable way to register the phenomenon for my thesis.

With regard to validity, the measure reflects the actual support for green parties rather inconsistently. The electoral system in a state may either favour or hinder representation of the green movement in the political sphere. Another problem is that this variable does not measure the amount of green seats relative to the actual amount of seats in parliament. The size of the legislature can vary dramatically between states, and if the amount of seats is supposed to be a proxy of the green party's political influence, the absolute number of seats does not tell us very much if parliament size is not included in the measure<sup>83</sup>. Despite the problems inherent in this measure, it is found valuable as such a variable is usually never applied on a global sample. Also, with clear theoretical expectations to this phenomenon, excluding it from the analysis would lead to a loss in the analysis.

The measure of the *left-libertarian executive* are obtained from the Database of Political Institutions, published by the World Bank (Keefer 2009). The variable *chief executive party orientation* is coded into left, centrist and right, based mainly on the content of the party name<sup>84</sup>. As the hypothesis indicates, my theoretical expectation is regarding *left-libertarian* leadership. To reflect this anticipation in my operationalisation, the variable is recoded into a dummy where left-wing and centrist executives are coupled together, with right wing executives making up the reference category. The coding of leftist and centrist governments into left-libertarian is based on the definitions provided in the codebook. Here, a left-wing party is defined as: "communist, socialist, social democratic, or left wing," whereas the centrist parties are defined as "centrist, or when party position can best be described as centrist (e.g. party advocates strengthening private enterprise in a social-liberal context)" (Keefer 2009:6). It should be noted that this interpretation of left, centre and right is based on

---

<sup>83</sup> To illustrate, the smallest country included in the sample, Bahrain, has a Chamber of Representatives holding 40 seats, whereas the most populous state, China, inhabits 2987 seats in the National Peoples Congress (CIA World Factbook 2010b). Thus, occupying ten seats in the Bahraini parliament would be dramatically different from occupying ten seats in the Chinese legislature.

<sup>84</sup> In the cases where party names did not indicate any orientation, several alternative approaches were used to identify party orientation. For a closer look at their other measurement mechanisms, see Keefer (2009:6-7).

a European understanding. Outside the developed world, Keefer (2009) has consulted additional sources to make sure that the party names actually reflected their ideological orientation. Where this distinction did not fit the categorisation, it has been coded 0. As such, I argue that the validity of this measure is satisfactory.

Lastly, *the regional effect* of IEA compliance is here understood as the percentage of countries in a region<sup>85</sup> that has complied at time  $t$ . As the fractional polynomials test revealed a logged relationship with the log hazard, the variable is logged in the upcoming analysis.

---

<sup>85</sup> The regions are taken from the United Nations (2000).



## **6. Analysing the determinants of compliance with the Montreal Protocol**

The aim of this thesis is to analyse what makes states comply with soft international environmental law within the frames of the Montreal Protocol. To do this, the biggest sample of member states possible has been collected, analysing behaviour from the onset of each member state's ratification. As this is one of the first statistical studies on party behaviour in IEAs, the main achievement of this thesis is to test different hypotheses based on previous research conducted on environmental performance, and apply it to the study of state behaviour within the study on compliance.

A Cox semi-parametric regression model will be used to look into the theorised expectations of the influence of national structural factors on the likelihood of compliance. The main assumption underlying this research tool is that the propensity to experience compliance either increases or decreases as covariates are included in the analysis. As this is a longitudinal study, I am able both to say something about the hazard at one point in time and to draw conclusions based on how members behave over time, thus strengthening our ability to make causal inferences.

This chapter starts off with introducing descriptive statistics of the variables. Next, bivariate Cox models are examined, revealing the relationship between the dependent variable and its covariates. Subsequently, multivariate Cox models will be introduced, allowing for statistical control. Based on the findings in the analyses, a synthesis model is established, followed by a thorough discussion related to the theoretical expectations.

### ***6.1 Descriptive statistics – range and distribution***

Before the results of the analysis are presented, a look at the distribution of data is important to create a proper understanding of the data used. As event history analysis does not assume normal distributed functions, skewed variables are not considered a substantial problem for the analysis (Cleves et al. 2008:2). Table 3 gives an overview of the minimum and maximum value, mean, standard deviation and observed N for the dependent and the independent variables, the latter being organised based on their theoretical affiliation.

**Table 3: Descriptive statistics**

	Variable type	Min.	Max.	Mean	Std.Dev	Obs N.
Dependent variable						
Compliance	Dummy	0	1	,24	,42	2513
Independent variables						
Realist factors						
Country strength (pop.)	Numeric	0,49	1314,36	43,31	139,99	2454
Ec. integration	Numeric	0,7	74,30	6,53	4,23	2380
Institutionalist factors						
Government autonomy	Numeric	1	18	2,88	1,74	2404
Government capacity	Numeric	25,4	97,1	70,72	17,83	2152
Corruption	Numeric	-2,09	2,39	-,15	,94	1845
Democracy	Numeric	-10	10	3,08	6,61	2419
Ec. development	Numeric	-1,83	4,20	1,56	1,29	2386
Common law	Dummy	0	1	,30	,46	2457
French civil law	Dummy	0	1	,44	,50	2457
Socialist law	Dummy	0	1	,21	,41	2457
German civil law	Dummy	0	1	,03	,18	2457
Scandinavian civil law	Dummy	0	1	,02	,14	2457
Constructivist factors						
IEA-membership(length)	Numeric	1	20	9,01	5,24	2457
EU-membership(length)	Numeric	0	57	2,02	8,90	2457
NGO activity	Numeric	,01	35,51	9,43	8,55	2390
Green Party Popularity	Numeric	0	55	1,05	4,99	2439
Left-libertarian executive	Dummy	0	1	,61	,49	1405
Regional compliance	Numeric	-5,50	2,10	-2,96	3,04	2015

The interacted variables are not included here as nothing substantial about the nature of the variable can be interpreted from these.

### 6.1.1 The dependent variable

When evaluating a binary variable, the interpretation of the mean and standard error values differs from that of a continuous variable. However, the mean gives information about the proportion of subjects attaining the value 1 in the variable (Midtbø 2007:44). Thus, for the dependent variable in this study, *compliance*, it is clear that 24 percent of the country-years<sup>86</sup> in this analysis attained 1 and consequently is in compliance.

<sup>86</sup> The designation *country-years* is used synonymously with observations in this analysis. It refers to the fact that there are several observations across time for each country, and thus that each observation is valid only for the country for a specified year.

### 6.1.2 Enforcement determinants

*Country strength* shows that the countries included in the analysis range between a population of 500 000 and 1,3 billion. Taking the mean into account, it is clear that most countries included in the study belong to the lower percentiles of this measure, identifying a skewed distribution in the predictor. *Economic integration* varies between 0,7 and 74,3, with a mean and standard deviation of 6,53 and 4,23 respectively. This tells us that most country-years find themselves inhabiting values between 2 and 11, and that few inhabit extreme values.

### 6.1.3 Managerial determinants

Government autonomy, government capacity and corruption, as well as democracy, are all numerical predictors. Even though the *government autonomy* variable ranges from 1 to 18, the mean of 2,93 tells us that the maximum value is an extreme, and that few veto players is a much more common situation than many veto players. *Government capacity* is measured through the Human Development Index which has been stretched out to range from 0 to 100. As the HDI is a composite index, it is hard to interpret exactly what a change in values means. Its mean reveals that most countries are placed on the upper half of the index, and this signifies that they are better at human development. However, it is hard to evaluate the distance between each value in this measure. *Corruption* ranges from -2,09 to 2,42 and has a mean of -,15 and standard deviation of ,94. The mean moves close to the median, signifying that most values are centred around the median values indicating an even distribution. However, the extreme values show that there exist states that deviate from this pattern. Whereas the *democracy* variable ranges from -10 to 10, the mean of 3,21 and the standard deviation value of 6,60 tells us that the states under scrutiny are quite evenly placed on the values, but that there is a majority of countries placed on the democratic side in the period under analysis.

As mentioned in the operationalisation, *economic development* has been logged, meaning its values have been stretched at lower values and compressed at higher values. This hardens descriptive interpretation as the distance between values is unknown. However, one can say something about the distance between values *per se*. This is important as the hazard rates in the upcoming analysis are based on a one unit change in the corresponding variable. Minimum value is -1,83 and reported maximum value is 4,20. This places the values more evenly out, which is visible through the mean value of 1,56 and a standard deviation of 1,29.

*Legal origin* has been divided into four binary variables, a fifth one serving as a reference category. First, the mean of Common law signifies that 30 percent of the member states have this legal system. Second, the mean of French civil law indicates that 44 percent of the member states have a French legal system. Fourth, member states practising a socialist law system make up 21 percent, whereas only three percent have adopted a German civil law system. The reference category, the Scandinavian civil system, makes up the two remaining percents of states in this study

#### **6.1.4 Legitimacy determinants**

The *IEA-membership - length* variable has values from 1 to 20, with 9,03 being the mean value and 5,25 the covariate's standard deviation. From this we understand that most member states are members to the Montreal Protocol from five to 15 years of the period under analysis. This also corresponds with the historical assessment of the Protocol. *EU-membership – length* has a skewed distribution. It ranges from 0 to 57, with a mean of 2,02 and a standard deviation of 8,90. Thus, the majority of countries in the analysis are given values at the lower end of the scale. Such a distribution is not surprising, taking into account that there are 154 countries in the analysis and 27 members of the EU. Moreover, the proxy variable for *ENGOS* ranges from ,01 to 35,51 with a mean value and standard deviation of 9,43 and 8,55 respectively. Thus, we see that also for this variable the distribution is not entirely even across the range of values, lower values being overrepresented in the sample.

The *Green party* covariate lets us know that even though the variable's maximum value is 55, this is an extreme case, as the mean and standard deviation when all county-years are considered are 1,23 and 5,28 respectively. This skewed distribution reveals what the theoretical discussion already suggested; that green party popularity is mainly a regional phenomenon in the western world, and that all in all, green parties tend to be quite marginal when seat representation is considered. *Left-libertarian executive* is a binary predictor. Based on the mean value, we find that in 60 percent of the subject-years in this analysis, a left-libertarian government is in charge. This, however, goes against my historical knowledge of the world between 1989 and 2008. Considering country-years, we find the reason for this unlikely finding. This is the predictor with the lowest N, and when looking at the data it becomes clear that the reported missing values are mainly seen in the developing world. This bias is unfortunate, and weakens the robustness of the findings for this predictor.

Lastly, *regional compliance* is also logged, rendering interpretation hard. Ranging from -5,50 to 2,10, but with a mean of -2.96 and standard deviation of 3,04, the predictor has been manipulated into distributing the country-year values evenly around the mean by stretching out lower values and compressing high values.

## 6.2 Bivariate Cox regressions

Bivariate regressions produce the first indications of the strength of the theories presented in chapter three. Here the independent variables are measured against the dependent variable separately, meaning the effect of other covariates on the single variable is not taken into consideration. However, bivariate regressions are important because they say something about how the covariates behave separately. Also, it reveals if the explanatory variables have any effect when regressed against the dependent variable alone, and if they do, if it is in the expected direction.

**Table 4: Bivariate Cox regression of compliance with the Montreal Protocol**

Independent variables	Hazard Ratio	Robust Standard error	Z	Number of subjects	Events	Time at risk
Country strength	1.001	.001	2.54	154	115	2031
Country strength inv.sq.	1.020	.011	1.91	154	115	2031
Ec. integration	1.021	.032	.66	150	114	1959
Government autonomy	1.039	.031	1.28	152	114	1982
Government capacity	1.001	.006	0.06	146	109	1742
Corruption	1.100	.131	0.80	146	108	1425
Democracy	1.038*	.017	2.39	153	115	1996
Ec. development logged	1.001	.076	0.01	150	114	1965
Common law	1.266	.268	1.11	154	115	2034
French civil law	.797	.163	-1.11	154	115	2034
Socialist law	.918	.210	-0.38	154	115	2034
German civil law	1.423	1.019	0.49	154	115	2034
Scandinavian civil law	9.473***	6.747	3.16	154	115	2034
MP membership – length	1.004	.026	0.16	154	115	2034
MP membership – length inv.sq	.976*	.012	-2.04	154	115	2034
EU membership - length	.987	.013	-1.02	154	115	2034
ENGO activity	1.006	.016	0.36	147	113	1969
Green party popularity	.976	.075	-0.32	153	114	2018
Green party pop*analysis time	1.005	.007	0.71	153	114	2018
Left-libertarian executive	1.021	.248	0.08	108	67	1121
Regional compliance logged	2.349***	.463	4.34	154	115	2034
Wald test Green pop*t, chi2	1.29					
Prob>chi2,	0.524					
Wald test MPlength(sq.), chi2	4.78					
Prob>chi2	0.091					
Wald test C.strength(sq.), chi2	6.72					
Prob>chi2	0.035					

Standard errors are robust. Efron method is used to handle ties.

Significance: \*\*\* p<.001 \*\* p<.01 \* p<.05 • p<.10. The tests are all two-tailed.

Rather few covariates come out of the model with a significant effect on the hazard of compliance. Starting out with the variables focusing on autonomy when explaining compliance, *integration in the world economy* is not found to be significant, confirming the null hypothesis of no effect. However, the inversely squared *country strength* reaches joint significance through the reported Wald test (see lower part of the table)<sup>87</sup>.

Among the variables focusing on the importance of an effective institutional framework, most factors have a positive effect on the likelihood of reaching compliance, corresponding with the theoretical expectations. The better *government capacity*, the better control on *corruption*, the more democratic and the higher *economic development*, the higher the hazard for experiencing compliance, corresponding with theoretical expectations. Nonetheless, only democracy reports a significant effect. The variable *government autonomy* shows that an increase in the number of veto players is followed by an increase in the hazard of compliance. Even though the result is not significant, this contradicts the theory assuming that the more veto players, the less autonomy and, consequently, less efficiency in the policy making process. Lastly, the dummy variables, making up the expectation of a countries legal system on IEA compliance, show mixed effects. *French civil law* and *Socialist law* both show the expected negative relationship, but fails to reach significance. *Common law* and *German civil law* show a positive effect, but fail to turn out significant, whereas the effect of *Scandinavian civil law* is extreme, reporting a significant nine times higher propensity for experiencing compliance if this legal system is at the base of a society. Two factors weaken the reliability of this finding. First, the high standard error reports that the variable is not to be trusted. Secondly, when inspecting the data, we find that only three countries in the analysis have this legal system<sup>88</sup>, and two of them comply early, explaining the extreme hazard reported. Thus, this finding should not be rendered much explanatory power.

Moving on to the factors emphasising social learning and legitimacy, mixed results are reported. The inversely squared *MP membership* function is through a Wald test found to significantly affect the chance of compliance. Moreover, *active ENGOS*, *left-libertarian executive* and *regional compliance* are found to have a positive effect on the likelihood of experiencing compliance. Only the last predictor is significant, though. Lastly, the *interacted green party popularity* fails to reach significance, and thus the effect of no relationship between the variable and compliance with the Montreal Protocol is upheld.

---

<sup>87</sup> The actual effect of the interacted and logged variables is only reported for the upcoming multivariate regression as it is here the main focus of the analysis lie.

<sup>88</sup> The countries are Norway, Sweden and Finland. Denmark and Iceland would also have been included here, but as mentioned earlier they are not included in the sample.

This initial analysis shows possible relationships in the data. However, the robustness of these findings is limited as they do not control for the effect of other variables, and consequently may be reporting spurious relationships in the data. Neither do they take into consideration the problem of collinearity between independent variables that is so common in social sciences analyses. In the following section, multivariate Cox regression models will be introduced, providing more valid results of the relationship between the likelihood of reaching compliance and its determinants.

### **6.3 Multivariate Cox regressions**

As already mentioned, a great advantage with multivariate regression is that it enables estimation of the real effect of the independent variables on the achievement of compliance with the Montreal Protocol, statistically controlled for the effect of the other covariates. Yet, including several covariates in the same regression creates new challenges to secure valid measurement of their separate effects. Thus, a premise for multivariate regression is that the effect of the explanatory variable (X) on the dependent variable (Y) is independent from the effect of the other covariates included in the analysis. If there exists correlation between two explanatory variables, *collinearity* is present in the data. This means that the effect of the two variables is hard to isolate as they tend to appear simultaneously for most subjects. As a result, accurate estimation of the relative contribution of each variable to the analysis becomes difficult (Skog 2004:286-288).

To get an overview of the presence of collinearity among the independent variables, a correlation analysis was conducted. Table 13 in Appendix reports the results<sup>89</sup>. The pairwise correlations show that the covariates *economic development* and *government capacity* have severe collinearity problems with a Pearson's R value of 0.92. This does not come as a surprise as HDI is used to measure government capacity, and this composite measure includes economic development as part of the index. Based on the theoretical interest in these two measures, they will both be included in the analysis, but will not appear jointly in the same model. Furthermore, some variables correlate at a problematically high level. This is the case for *corruption*, which shows values above 0.7 when paired with *ENGO-activity* and *economic development*. Economic development and ENGO-activity also correlate on this level. This

---

<sup>89</sup> For Pearson's R scores below 0.5, collinearity is considered modest, with limited influence on the results. Values of 0.6 and 0.7 lead to big standard errors, and as the values come close to 0.8 the data becomes unreliable (Skog 2004:288).

indicates that the three covariates tend to measure similar phenomena as they appear together most of the time. Because the corruption predictor is weakly justified and is stricken by missing data, it is excluded from the analysis<sup>90</sup>. Lastly, the pairwise correlations reveal that the stratification variable, Article 5-membership, shows Pearson's R-values around 0.6 with five covariates (Government capacity, NGO-activity, Regional Compliance, Economic development and Corruption). When stratifying the models after this variable, then, we risk that parts of the covariates' effect is absorbed by this variable. Still, the theoretical justifications for choosing a stratified model is more important than maximising model fit, I argue.

### 6.3.1 Compliance with the Montreal Protocol – the results

Table 6 reports the results of the effect country-specific factors have on the chance of reaching compliance with the Montreal Protocol. The composition of the models has been chosen based on the dual wish to maximise N as well as including as many relevant variables as possible. Thus, the first model has the largest N and least covariates included. As more covariates are enclosed in the analysis, the amount of observations drops due to missing data.

Recall that the results reported are *hazard ratios*, not coefficients. This means that they must be interpreted as relative to 1. Reported values below 1.0 signify a negative effect on the hazard rate per unit positive change in the covariate. On the opposite side, values above 1.0 indicate a positive effect between a one unit increase in the explanatory variable and the hazard ratio for reaching compliance. To exemplify, a democracy variable reporting a hazard ratio of .95 would mean that one unit change in the democracy covariate leads to a five percentage point reduction in the hazard rate (meaning it is less likely that compliance is reached relative to the previous value). On the other hand, a hazard ratio of 1.05 tells us that for a one unit increase in the democracy variable, the chance of experiencing compliance heightens with five percentage points compared to the previous covariate value. Also, one should keep in mind that the Cox regression performed is stratified, meaning that *all hazards reported are controlled for inter-member differences* with regard to the postponement on emission reduction given to Article-5 countries.

---

<sup>90</sup> When the analysis is run including the corruption variable, it is far from statistically significant, also when the correlating covariates are excluded from the model. This is a strong indication that the corruption covariate does not affect the increase in the hazard ratio for experiencing compliance. Furthermore, a Wald test is also run, reporting a chi2 value of 0.33, thus showing that leaving the Corruption variable out does not lead to a significant reduction in model fit. Consequently, we can be quite certain that leaving this variable out of the analysis does not lead to decreasing explanatory models.



The Cox model assesses coefficient significance using Wald statistics. Thus, the reported p-values are based on reported Z scores<sup>91</sup>. Tests of model fit are introduced at the bottom of table 6. As mentioned earlier, the log-pseudolikelihood test and Wald chi squared test report whether the models significantly contribute to explaining the nature of compliance measured against an empty model. The higher value, the better the covariates are at explaining the phenomenon under scrutiny. When comparing across models, only the Wald chi squared test is used due to reduced robustness of the -log pseudolikelihood measure. Related is another model fit test, linktest, which reports whether the model ( $\hat{\mu}$ ) is correctly specified as opposed to an alternative model ( $\hat{\mu}_{sq}$ ) with random covariates<sup>92</sup>. Lastly, univariate Wald tests for the interacted variables are included for all three models; showing that in all models the inclusion of the original variable alongside the interacted covariate significantly improves the model. All models cover the period 1989-2008. Considering the limited number of country-years, especially in model 3, a .10-level will serve as the cut-off point for significant results. Two-tailed tests are run on all variables.

---

<sup>91</sup> These are interpreted the same way as the  $t$  values used for testing coefficient's significance in linear regression models (Hair et al. 2006:273).

<sup>92</sup> The linktest is a test verifying that the specifications of covariates are good by searching for hypothetical variables to add to the model. The variable  $\hat{\mu}$  should be significant, whereas the variable  $\hat{\mu}_{sq}$  should be insignificant. This is because  $\hat{\mu}$  represents the predicted value from the model.  $\hat{\mu}_{sq}$ , on the other hand, should not have much predictive power except by chance if the model is properly specified.

**Table 5: Multivariate analysis of the Montreal Protocol**

<b>Independent variables</b>	<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>
Country strength	1.001** (.000)	1.000 (.000)	.999 (.001)
Country strength inversely squared	1.034** (.013)	1.044** (.017)	1.046* (.057)
Integration in world economy		1.006 (.043)	.980 (.040)
Government autonomy		.961 (.048)	1.061 (.050)
Government capacity			.989 (.017)
Democracy		1.022 (.022)	.976 (.032)
Economic development		.818* (.079)	
Common law	.997 (.832)	1.061 (.934)	1.450 (1.437)
French civil law	1.199 (1.037)	1.256 (1.150)	1.090 (1.175)
Socialist law	.859 (.764)	1.126 (1.105)	1.593 (1.847)
German civil law	.981 (.783)	1.002 (.810)	1.058 (1.209)
MP-membership (length)	1.016 (.031)	1.024 (.037)	1.046 (.057)
MP-membership inversely squared	.967** (.012)	.958** (.016)	.951* (.022)
EU-membership (length)	.966* (.016)	.963• (.020)	.968 (.022)
ENGO-activity		1.029 (.038)	1.014 (.060)
Green party popularity	.905 (.076)	.902 (.082)	.910• (.051)
Green party pop*analysis time	1.016* (.006)	1.016* (.007)	1.018*** (.004)
Left-libertarian executive			1.201 (.400)
Regional compliance	2.633*** (.554)	2.507*** (.506)	2.424*** (.521)
Number of country-years	2015	1855	977
Number of compliance events	114	111	66
Log-pseudolikelihood	-397.18	-374.42	-169.18
Wald chi2	44.42	49.19	76.01
Prob > chi2	0.000	0.000	0.000
Linktest, _hat	0.000	0.000	0.000
Linktest, _hatsq	0.539	0.456	0.424
Wald test Green pop.*t, chi2	9.99	8.92	29.99
Prob>chi2, Wald test Green pop.*t	.007	.012	.000
Wald test, MP-length (sq.), chi2	8.60	9.00	6.84
Prob>chi2, Wald test MPLength(sq.)	.013	.011	.033
Wald test, Country strength (sq.)	9.56	10.84	4.81
Prob>chi2, Wald test C.strength(sq.)	.008	.004	.092

Reported hazard ratios. Robust standard errors in parentheses. Efron method is used to handle ties.  
Significance: \*\*\* p<.001 \*\* p<.01 \* p<.05 • p<.10.

### 6.3.2 Model one

Model one has the biggest sample, with 2015 country-years<sup>93</sup> and 114 instances of compliance occurring. Recall that the Cox regressions are all stratified, and that the reported results takes into account the heterogeneity found between the two different groups of members to the Montreal Protocol. Thus, significant results see the same patterns in both member camps.

*Country strength* was identified as having a non-linear trend through fractional polynomials and was thus included alongside an inversely squared version in the regression. As with all squared functions, *country strength* and *country strength inversely squared* can not be interpreted independently. Even though both predictors turn out statistically significant in model one, the univariate Wald chi squared test shows that jointly, the covariates are significant at the 01.-level. To interpret the predictor's effect on the hazard ratio, an equation was run on the coefficients. The results reported in Table 16 in Appendix show a monotonic negative effect of a one million increase in population at different values of the predictor. There is a steep and clear trend towards 50 million, and after that the effect of one million more inhabitants is diminishing. After 100 million inhabitants the hazard ratios decrease has levelled up around 0, meaning that a one million rise in population reduces the likelihood of moving into compliance with the Montreal Protocol by 100 percentage points<sup>94</sup>.

None of the dummy variables included to measure the effect of legal system turn out statistically significant. As such, they confirm the null hypothesis of no connection between legal system characteristics and the chance of reaching compliance with the Montreal Protocol. Considering the theorised relation, *Socialist law* shows the expected negative sign when moving from *Scandinavian law* (the reference category), whereas *French civil law* shows a positive effect on the hazard of experiencing compliance, contradicting its theoretical basis. *Common law* and *German civil law* were also found to be insignificant. Recall that they were included to create flexibility in the subject of interest, and thus are not of theoretical

---

<sup>93</sup> Originally, the sample was considerably bigger, containing 2457 country-years. However, when telling STATA that my data are to be used in an event history analysis (using the `stset`-function), STATA excludes 423 observations because they include information about the subjects after they have reached compliance, thus exiting the analysis. The final sample then, is on 2034 observations.

<sup>94</sup> One should however be aware that there are only ten countries in the sample with a population of more than 100 million, meaning that the finding may be an artefact of the limited number of observations at the high population level as well as of the squared equation used in the estimation. Still, running the equation on a sample excluding the extreme outliers China, India and the US, shows the exact same curve, indicating that the trend seems to reflect the data.

importance. Big standard errors combined with non-significant findings render all legal system results unreliable<sup>95</sup>.

Like the country strength predictor, *MP membership - length* is evaluated through an inversely squared function. Running a Wald Chi squared test shows that the two predictors are jointly significant with a p-value of 0,014, thus their effects must be regarded as valid for the larger population of member states. Looking at its effect for a one year change on different values of the predictor, we find that the effect is positive across all values. Yet, the effect is more modest for a one year increase on smaller values before it starts increasing rapidly from ten year onwards. The first year a signatory state's chance of reaching compliance increases with 16,7 percentage points for one additional affiliate year. Yet, as states have been members for ten years, the hazard doubles for one year added and as one moves towards twenty the reported hazards become extreme<sup>96</sup>.

*EU-membership – length* is significant at a five percent level. The results indicate that a one extra year as an EU-member leads to a 4,4 percentage point reduction in the likelihood of reaching compliance. This supports the theories assuming better compliance records for new EU-members.

The *Green party popularity* failed to pass the Schoenfeld proportionality test. Consequently, an interacted version, multiplying Green party popularity with analysis time is included alongside the original variable. Like the squared functions, this interaction can not be evaluated without both variables being jointly included in the analysis. A univariate Wald chi squared test confirms that the two variables combined are significant and thus improve model fit. The direction of the connection signifies that the original Green party variable sees a negative development, whereas the interacted variable shows a positive effect on the hazard ratio. To be able to interpret the effect of this non-linear relationship, the covariates' coefficients are run through an equation reported in Table 14 in Appendix. The table reveals that the effect of the variable on the likelihood of experiencing compliance changes over time. Initially, the hazard ratio is negative, before the function changes direction after 6,4 years. After this, the effect slowly turns more positive, and for the last observation point the hazard has jumped to 1.219, indicating that one extra green seat in parliament in year 19 leads to 21,9

---

<sup>95</sup> The high standard errors can probably be explained by the few observations present on some of the dummies. Both Scandinavian law and German civil law has got way too few observations to see robust results here. Furthermore, bivariate variables tend to have higher standard errors than metric variables as there is smaller variation across the sample (Skog 2004:286).

<sup>96</sup> A probable reason for the extreme effect on the hazard ratio on higher values is that most countries in this analysis do actually comply with the Montreal Protocol during the analysis time, and this rate is especially high for the ones who have been members the longest. This should affect the hazard substantially.

percentage point bigger chance of fulfilling the compliance obligations to the Montreal Protocol.

*Regional compliance* is found to have a significant effect on the likelihood of reaching compliance at the .001-level. The reported positive effect is in accordance with the theoretical expectations. A one unit increase in the variable reports an extreme 163 percentage point raise in the hazard ratio of experiencing compliance. As this variable is logged, measuring absolute change does not give much intuitive meaning (Skog 2004:248). Thus, the reported coefficients are modified into measuring the effect of one percentage point change in the predictor<sup>97</sup>. When doing this, we find that for a one percentage point positive change in regional compliance the hazard ratio is raised by 2,7 percentage points. This may sound like a minimal effect, but keep in mind that the variable is continuous and contains 100 values, meaning that the difference between units is qualitatively smaller than the distance between units in a bivariate covariate.

In terms of overall fit, Model one performs well. The  $-\log$  pseudolikelihood value reports -397,18, which must be considered good on the premises that the higher the negative number, the better the model. Furthermore, the Wald chi squared test measuring the strength of the model is statistically significant at a .001-level, concluding that the model is better than the null model of mere chance. The test itself of 44,42 is not particularly high. As the test measures covariates' strength at explaining accord compliance, this is an indication that the model performs moderately when explaining the parameters. Lastly, linktest performs well, showing that Model one ( $\hat{\mu}$ ) includes well specified and significant covariates, whereas the competing model ( $\hat{\mu}_{sq}$ ) representing mere chance is far from being statistically significant.

### 6.3.3 Model two

Model two includes five new covariates into the analysis, consequently reducing the number of observations and compliance incidents to 1855 and 111, respectively. *Integration in world economy* fails to meet statistical significance, and furthermore seems to have close to no effect on the likelihood to reach compliance with a hazard ratio of 1.006. Despite its minimal effect, the result is in the hypothesised direction. The negative effect of *government autonomy* is in accordance with theoretical expectations, but fails to significantly affect a member state's compliance record. Moving on to *democracy*, a positive relationship is found, corresponding

---

<sup>97</sup> To find the effect of a one percent rise in a logged covariate, the coefficient is divided by 100. Furthermore, the hazard ratio is detected by exponentiating the modified coefficient.

with theory, but again a p-value of .292 shows that the results are not statistically robust. The heavily theorised *ENGO-activity* is no better than the three previous predictors, reporting a positive relationship that is not verified because of lack of statistical significance. In contrast to the four abovementioned covariates, *economic development* shows a significant, negative relationship with the dependent variable. Recall, however, that the variable is logged. When converting the coefficients, I find that for a one percentage point increase in the variable, the hazard of reaching compliance increases by 0,8 percentage points<sup>98</sup>. This means that the richer the country, the more likely it is to fulfil the requirements set out in the Montreal Protocol. Confronting the results with theory, we see that this finding corresponds with the well established assumption that richer countries are better compliers.

The introduction of five new variables only leads to small changes in reported effects in the other variables. First, a Wald test of the joint contribution of the original and the interacted Green party popularity variables is now only significant at the .05-level. Still, the effect of the variable on compliance is almost identical to the one found in Model one. Secondly, the reported effects of the *Common law*, *Socialist law* and *German civil code* dummies have all changed direction. Now both Socialist law and French civil law show a positive effect on the chance of reaching compliance, contradicting the initial assumption. As already mentioned, problems with the measurement itself as well as the lack of significance may account for the big changes in reported effects.

Considering model fit, Model two performs similarly to Model one. Despite the increase in predictors included in the regression, the multivariable Wald chi squared test shows a small increase from Model one to 49.19. This indicates that the two models to a limited extent, although significantly, can explain the parameters under scrutiny. The -log pseudolikelihood is reduced, but not considerably. The linktest reports that the variables included are of importance and that their specification is good.

### 6.3.4 Model three

Model three is built on the smallest sample, including 977 observations and 66 instances of compliance occurring. The reason for the substantial decline is that two variables with a considerable amount of missing data are included in the model. These are the *Government capacity* and *Left-libertarian executive* predictors<sup>99</sup>. Both of the new covariates are expected

---

<sup>98</sup> Because  $0,818/100 = 0,00818$ , and  $\exp(0,00818) = 0,8\%$ .

<sup>99</sup> As the *Government capacity* and *economic development* measures show problematic instances of collinearity, the former is removed from this model so the actual effect of the latter covariate can be measured.

to vary positively with the likelihood of reaching compliance. *Government capacity* has a negative effect on the chance of compliance, running counter to theoretical postulations. Nonetheless, a p-value of .505 upholds the null hypothesis of no effect between *Government capacity* and the likelihood of protocol compliance. The positive effect of *left-libertarian executive* is in accordance with theory, but also this predictor fails to reach a minimum of significance, confirming the null hypothesis of no relationship.

The effects of most variables remain similar to the previous model. Nonetheless, some exceptions are present and should be mentioned. First, the effect of *EU-membership* fails to reach statistical significance. Secondly, the significance level of *country strength* decreases, barely reaching a .10-level joint effect, whereas *Green party popularity* turns out significant at the .001-level. Thirdly, the same *green party* predictor changes direction earlier than what was reported in the two previous models. Fourthly, several predictors report a change in effect direction. *Government autonomy* moves from being negative in Model two to positive here, and *integration in world economy* as well as *democracy* reports of negative effects as opposed to the positive effects seen in Model two. The change in direction is likely to be caused by the substantial decline in country-years seen in Model three. Still, all these covariates are statistically insignificant.

When evaluating model fit, Model three reports a Wald chi squared value of 76,01, significantly better than the previous two models. This is an indication that it reflects the world more properly. Missing data is a problem here, following the two included variables. As reported in the descriptive analysis, *left-libertarian executive* systematically misses values from the developing world. Knowing this, a possible reason why Model three outperforms the two former might be that mainly developed countries are included, and the variables included in my model fits the developed world better than the developing world. The linktest is still significant for the model under question and insignificant for the model of mere chance, indicating that the correct covariates are included and that they are soundly specified.

#### **6.4 Constructing a synthesis model**

Wishing to present the best explanatory model of compliance with the Montreal Protocol, the synthesis model will include all variables that have turned out statistically significant in the three former models, regardless of their theoretical affiliation. The covariates meeting this criterion are summarized in the table below.

**Table 6: Overview of variables with significant effects in Models 1, 2 and 3**

	Variable	Model 1		Model 2		Model 3	
		Sign.	Direction	Sign	Direction	Sign.	Direction
H1/H2	Country strength	Yes	Negative	Yes	Negative	Yes	Negative
H10	Economic Development	-	-	Yes	Negative	-	-
H13	MP-membership	Yes	Positive	Yes	Positive	Yes	Positive
H15/H16	EU-membership	Yes	Negative	Yes	Negative	No	Negative
H18	Green party popularity	Yes	Neg/Pos	Yes	Neg/Pos	Yes	Neg/Pos
H20	Regional compliance	Yes	Positive	Yes	Positive	Yes	Positive

The cells where Neg/Pos is written indicate that this is an interacted variable where the effect changes from negative to positive over time.

Table 7 includes six different covariates, as well as two inversely squared functions of *Green Party popularity* and *MP membership – length* which are still found to increase model fit significantly when included next to the original first degree variables. When checking for the proportionality function in a Schoenfeld test, we find that *Green party popularity* does no longer break with the proportionality assumption, meaning the interaction with analysis time is not necessary for the synthesis model<sup>100</sup>. The significant variables were already included together in Model two, but then many inefficient predictors were also included, possibly influencing the effect on the other predictors as well as lowering model fit.

<sup>100</sup> See the Schoenfeld results in Table 11 in Appendix



**Table 7: Synthesis models**

Independent variables	First Synthesis model	Final synthesis model
Country strength	1.000* (.000)	1.000 (.000)
Country strength inversely squared	1.041** (.014)	1.043*** (.013)
Economic development	.839• (.076)	.818* (.073)
MP-membership	1.041 (.032)	1.042 (.032)
MP-membership inversely squared	.967** (.011)	.968** (.011)
EU-membership	.981 (.018)	-
Green party popularity	1.029 (.029)	-
Regional compliance	2.437*** (.480)	2.452*** (.481)
Number of country-years	1947	1947
Number of compliance events	113	113
-Log Pseudolikelihood	-391.230	-392.373
Wald chi2	38.06	36.62
Prob>chi2	0.000	0.000
Linktest, _hat	0.000	0.000
Linktest, _hatsq	0.292	0.247
Wald test, MP-member(squared), chi2	11.78	11.88
Prob>chi2, Wald test MP-member(sq.)	0.003	0.003
Wald test, Country strength(squared), chi2	9.67	11.02
Prob>chi2, Wald test C. strength(squared)	0.008	0.004

Reported hazard ratios. Robust standard errors in parentheses. Efron method is used to handle ties.

Significance: \*\*\* p<.001 \*\* p<.01 \* p<.05 • p<.10.

Looking at the first synthesis model, the statistical merit is high. Most of the covariates show statistically significant effects that correspond to the theoretical postulations. The less powerful a state is, the more prosperous it is, the longer it has been member to the Protocol and the more compliant region it belongs to, the higher propensity for complying with the Montreal Protocol. Yet, two of the variables, *Green party popularity* and *EU-membership*, are rendered statistically insignificant with p-values of .298 and .288 respectively<sup>101</sup>. This result might not be that surprising for *green party popularity*, as it has been included alongside an interaction term in the previous models because of non-proportionality. It seems that when the variable is not interacted with time, it fails to reach significance, and thus the null hypothesis of no connection between green party success and the likelihood of reaching compliance stands. *EU-membership - length* also reports results that make their inclusion in a final

<sup>101</sup> The MP-protocol covariate also fails to reach significant results. Yet, when coupled with the transformed variable it reports significant results and is thus kept in the model.

synthesis model difficult. However, we already saw that this predictor failed to reach significance in Model three, indicating that the connection was not that strong. Based on their insignificant results of the two variables, the hypotheses H5 and H7 are not supported by the model: the success of *green parties* as well as the length of *EU-membership* in a state does not significantly affect the chances of reaching the goals put out by the Ozone Secretariat.

In terms of model fit, we see that the first synthesis model performs moderately. The multivariate Wald chi squared test is quite low, but significant. This means that the model helps explaining the issue of compliance substantially, but that the strength of the model in explaining the parameters is reduced. The -log pseudolikelihood reports high values, indicating good model fit, but as this can not be compared across models it does not tell us much. The linktest tells me that I am still including the correct predictors, and that their specifications are sound.

Based on the above-mentioned considerations, I conclude that the model can still be improved. Thus, wishing to maximise efficacy, parsimony and model specification, a final model is presented excluding the two insignificant variables from the first synthesis model. To detect whether the removal of the two variables can be justified, a univariate Wald test chi squared is conducted to see whether they improve model fit significantly.

**Table 8: Wald test of the Green popularity and EU-length covariates' joint effect**

Wald test	
Chi2(2)	1.19
Prob>chi2	.553

The far from insignificant p-value tells us that the null hypothesis asserting that the variables do not influence the model significantly should be upheld. Thus, removing Green party popularity and EU-membership will not influence the model's explanatory power significantly.

The final synthesis model is reported in the right column of Table 7. In terms of variable effect, the model largely corresponds to the first synthesis model. The covariates have statistically significant effects in the same directions, and their reported Z-scores have a close to identical magnitude. *Economic development* is now significant at the .05-level with a p-value of .024 in the final model. The statistical effect has also improved for the *Country strength* interaction term. Considering model fit, the two synthesis models are close to identical.

## 6.5 Discussion

The examination of the models has confirmed that the results reported in Table 5 and Table 7 perform relatively well and that their results are robust. This section will discuss the results of the analysis more in detail and connect them with the hypotheses. When assessing the size of predictor effects, the model with the largest number of observations and cases of compliance available for the predictor are chosen to maximise reporting soundness.

*Country strength* shows a large, negative effect over time. Based on Table 15, an increase from one million to two million inhabitants decreases the chance of experiencing compliance with the Montreal Protocol with 11,8 percentage points. For a state with 30 million inhabitants, a one million increase would decrease the hazard with 92,1 percentage points relative to the baseline hazard. To exemplify these findings, one can picture Algeria who had 30 million inhabitants in 1999, but reached 31 million in 2001. From 1999 to 2001 then, the likelihood of Algeria complying with MP emission standards decreased by 92,1 percentage points. Zambia, on the other hand, reduced its chances of reaching compliance with ‘only’ 58,3 percentage points between 1999 and 2004 as their number of inhabitants increased from 10,2 to 11,2 million. Hence, I conclude that the more powerful the country, measured in the shape of its population, the less likely it is to adhere to the Montreal accord. As mentioned earlier, we see a stabilising effect around 0 from 90 million onwards, and from there the effect of one extra million inhabitant does not have an additional decreasing effect on the hazard.

Confronting the results with theory, my analysis of compliance with the Montreal Protocol finds that the less country strength, the more likely a member state is to move into compliance, regardless of their Article 5-status. Thus I can conclude that out of the two hypotheses presented regarding country strength; only H1 is supported. This finding gives strength to Perkins and Neumayer’s (2007) assumption that powerful states legitimate themselves through their importance economically and politically and thus do not depend on cooperation within international agreements to be considered an actor in the international community<sup>102</sup>. This observation serves as support to the realist school arguing that states always evaluate their self-interest before moving into binding IEAs. Since powerful states can afford to act more autonomous internationally, they are less likely to join in on international efforts to save the environment if it is more costly than beneficial to them.

---

<sup>102</sup> Even though generalisations to other Protocols are not methodologically sound, I find a good analogy in the observed role of the US with regard to the Kyoto Protocol in the late 1990s, where the country was included in the negotiations of the regime framework, but in the end did not ratify the agreement.

Moving on to *economic integration*, this analysis allows no significant assumption about the relationship between level of integration and the propensity of experiencing compliance, thus rejecting both H3 and H4 theorising such a relationship. This supports Bernhagen's (2008) findings, even though his findings is of a much smaller scope than mine, including 35 advanced industrialised countries. Not only is the variable rendered insignificant, it also changes direction from Model two to Model three. This is most likely due to the considerable drop in number of observed country-years for Model three. Despite this change in direction, the variance in effect is minimal and both hazard ratios concentrate around 1, signifying no effect. The lack of effect is also further confirmed by the bivariate regression reporting that even when the effect is singled out, it fails to reach significance. I thus conclude that the level of trade dependency can not explain the observed variance in compliance with the Montreal Protocol.

The number of veto players in a political system is found to decrease the chances of experiencing compliance, thus corresponding with H5. However, the effect is not significant, and the effect changes direction as the regression is run on Model 3. The initial bivariate regression also showed a non-significant relationship between *government autonomy* and compliance. These findings thus contradict Haverland (2000) and Giuliani's (2003) findings of government autonomy's positive effect on environmental performance. Based on the models presented here, I confirm the null hypothesis assuming that the degree of government autonomy does not affect the likelihood of reaching compliance with the Montreal Protocol. Thus, H5 is dismissed.

As *government capacity* is only included in model 3, its effect should be interpreted with some care as there is a risk of sample bias. The regressions show a negative effect on compliance as HDI increases, dismissing the postulated positive effect expressed through H6. It also contradicts the positive effect found in the bivariate analysis. This finding challenge the assumption by Haas et al. (1993) and Simmons (1998) that it is crucial for IEAs to facilitate administrative and technical capacities to increase compliance records among member states. Based on the findings in this analysis, H6 is rejected.

*Corruption* was excluded from the analysis because of problematically high correlations with several other predictors. Hence, H7 is not answered in this thesis. This is a weakness, but the bias risked from including the covariate made its inclusion methodologically unsound.

Despite *democracy* coming out of the bivariate model as significant at the .05-level, this effect is lost as the predictor is included in the regression alongside other variables. Both

Models 2 and 3 fail to produce significant support for H8, hence the null hypothesis of no effect of the political regime on the likelihood of reaching compliance with the Montreal accord stands. This finding contradicts previous studies presented in the theory chapter, and thus one might suspect that the model specifications may have obstructed the data. Nonetheless, the descriptive statistics displayed variation and no signs of a biased sample. When running the regression on the sample of Model two without the variables that democracy correlates the most with, it still turns out as non-significant. As most of the previous empirical studies conducted have a focus on developed countries, their significant positive effect for democracy on compliance is not that surprising. As such, it might be assumed that the positive effect seen for the developed world is overshadowed in this analysis by the mixed record of the non-Article 5 members to the Protocol. Thus, I conclude that the political regime of a member state can not explain compliance with the Montreal Protocol.

The effect on the covariate *economic development* is negative and significant across models. Yet, as the natural logarithm is taken of this variable, a different interpretation is required. Doing this, I find that a one percentage point increase in the variable increases a state's chances of fulfilling its obligations by 0,8 percentage points (based on the final synthesis model). This may seem like a minimal effect, but one should keep in mind that this is a continuous variable, meaning that there are much smaller intervals seen here than for a dummy variable where a one unit increase is a move from the one extreme to the other. As such, I reject the null hypothesis and strengthen H10; the better the economic situation in a country, the more likely it is to reach the protocol emission standards and move into compliance. Applying some math, we find that the boom in the Malaysian economy from 1989 to 1996 increased the country's chances of complying with the Montreal Protocol with 6 percentage points.

Based on this assessment of economic development, I conclude that the findings of this thesis with regard to the importance of environmental development correspond with some of the literature on the field (Brown Weiss and Jacobson 1998; Jänicke 1992). The importance of available resources seems obvious when extensive changes are to be done in a society. This further is an argument for an increased focus on compliance funds to help the poorer states reduce emissions according to treaty goals. Such managerial help serves as an argument for the institutionalist school arguing that non-compliance is not a matter of recalcitrance but of inability. It should also be pointed out that this analysis does not find support for H9 arguing for the existence of an environmental Kuznets Curve. As mentioned previously, this squared effect is not found when other covariates are accounted for, thus supporting Neumayer (2003).

*Legal traditions* are theorised to influence the propensity to comply with the Montreal accord, but the legal origin dummies included in the analysis all fail to predict the outcome. Societies grounded on *French civil law* and *Socialist law* were both theorised to be negative for compliance, but in the analysis the former is reported positive against all models whereas the latter shows the expected effect in Model one before turning positive in the two latter. It should however be pointed out that the descriptive statistics made it very clear that the reference category make up two percent of the country-years included in the sample. This means that the effect of the other dummies hinges on how three Scandinavian states (Norway, Sweden and Finland) perform at every time. This can also explain the great fragility observed in the data with regards to which other covariates were included. Nonetheless, the results are far from significant, rejecting H11. Hence, there is no effect to be found, neither between the two theorised legal systems nor the other legal dummies included in the analysis.

H12, hypothesising a positive effect between public environmental concern and compliance, could not be tested in the analysis due to lack of comprehensive data from the WVS and EVS. This is a weakness, but the decision is in line with the methodological request of avoiding too much missing data and consequently risking sample bias. Thus, I am not able to conclude on whether the postulations made by Haas (1998) and Brown Weiss and Jacobson (1998) correspond with an empirical analysis of the Montreal regime.

The effect of *length of membership* to the Montreal Protocol turns out statistically significant and consistent with the theoretical expectations across all models, and thus represents one of the strongest findings in this study. I find that the effect of one additional affiliate year on the propensity for experiencing compliance is non-linear and takes a monotonous, curvilinear form in all models. Table 16 shows how the hazard ratio of experiencing compliance for one additional member year increases gradually the longer a state has been member to the Montreal Protocol. For the final synthesis model we see that as a new member, the first year will increase the hazard with 5,7 percentage points compared to the hazard for non-members. After five years as a member, an additional associate year increases the likelihood of complying with the accord with 36,7 percentage points relative to the baseline. To illustrate the effect found here we take a quick look at Switzerland. In 1993 they had been members to the Protocol for five years. In 1994 then, their chance of reaching compliance had increased with 36,7 percentage points in respect of the hazard in 1993. From ten years onwards the effect of an extra affiliate year starts growing rapidly, and a state that has been a member for 15 years improve their chances of moving into compliance with 160,2 percentage points for an additional Protocol year. These high numbers make sense when

taking into consideration that the average state needs 9,5 years to move into compliance with the Montreal accord. This means that for each additional year after this, there are fewer and fewer cases left in the sample, increasing the hazard substantially.

As the few studies that have discussed the effect of membership have been qualitative in their approach, the particular shape of its relationship with compliance has not previously been theorised. Thus, when this analysis finds a positive and significant curvilinear trend over time, it is the first statistical analysis to do so. The results confirm H13; the longer a state has been a member to the Protocol, the bigger the chance that it will move into compliance (controlling for the effect of the Article 5-delay). Connecting the findings to theory, there is a case to be made for the theorised effect of social learning and maturity in a compliance process. As affiliates interact with each other on this arena, they internalise the norms underlying the treaty and use diplomatic connections to compel other states into taking the compliance question seriously. In the case of the Montreal Protocol, there is no doubt that the supervisory mechanisms have played an important role in driving member states toward compliance. The sanctioning possibilities available to the Members of the Parties<sup>103</sup> as well as the availability of support for capacity building are likely to effectively steer for compliance. This, however, opens up for more rational mechanisms underlying the positive effect of the time variable. Maybe it is not internalised norms but rather a cost-benefit analysis by member states driving compliance with the accord? Even though I can not single which mechanisms explain compliance variance, the findings allows me to conclude in concordance with Brown Weiss and Jacobson (1998:512) that the trend over time is towards greater compliance, at least for the Montreal Protocol.

The examination of *EU-membership length* ended in three hypotheses in the theory chapter. Recall, however, that H14 was not tested in the analysis because of variable collinearity. The four models including the covariate all show a negative hazard, signifying that the longer a state has been member to the EU, the less likely he is to move into compliance. This supports H15, as well as Giuliani (2003) and Perkins and Neumayer (2007), arguing that new EU-members are more interested in showing international responsibility than their mature adversaries. Yet, the skewed distribution uncovered in the descriptive statistics might be part of the explanation for this finding. As there are only 27 out of 154 states under study that are EU-members, each EU-member failing to comply at time  $t$  counts disproportionately much, potentially skewing the results. Nonetheless, the relationship is

---

<sup>103</sup> The Member of the Parties have the right to suspend membership rights as well as stopping technical or financial transfers (Breitmeier et al. 2009).

significant in two of the original models, allowing us to reject both H16 as well as the null hypothesis of no connection. The fact that the predictor fails to reach a significant result in the first synthesis model seems to be connected to the exclusion of other covariates rather than the N as the model is almost as big as Model one where EU-length was found to be significant. Thus, it might be that the effect of EU-membership on the hazard of compliance is contingent upon other explanatory variables. Nonetheless, the conclusion is that there exists fairly robust evidence that young EU-members (and non-members) are better compliers than their predecessors.

That *ENGO activity* plays an important role as a compliance pusher is one of the most established theories presented in this thesis (Breitmeier et al. 2006; Brown Weiss and Jacobson 1998; Cameron et al. 1996; Checkel 2001; Haas et al. 1993). Thus, its lack of ability to reach significance in this thesis comes as a surprise. Despite showing the hypothesised positive relationship between the amount of active NGOs in a society and its compliance record, H17 is rejected in favour of H0 concluding that no effect exists. One possible explanation for this finding might be that the proxy variable measuring active NGOs fails to reflect the actual effect ENGO activity has on the likelihood of reaching compliance. As mentioned in the data chapter, this might not be a good measure. It does not automatically follow that an open society with many NGOs has an environmentally friendly population. Thus, I state that there is no connection between presence of ENGOs and protocol compliance in this thesis. Still, it is possible that with a better measure available, I could have gotten contrasting results.

As thoroughly covered, *Green party popularity* revealed non-proportionality in three models when run through the Schoenfeld test<sup>104</sup>. This indicates that the predictor effect varies with time. Thus, a covariate interacted with analysis time has been included alongside the original variables in the three models of Table 6. The two variables are jointly found to significantly influence the chance of experiencing compliance in all models. Table 14 shows the difference in hazard ratios for different values of the Green seats variable. Initially, one green seat leads to a nine percentage point reduction in the hazard of experiencing compliance compared to those member states not having any green seats. Nonetheless, this negative effect diminishes quickly, and after 7 years the effect of an extra green seat in parliament has turned positive. After 19 years (maximum amount of observation year a member state has is 20) one extra green seat increases the hazard of experiencing compliance by 22 percentage points.

---

<sup>104</sup> See Table 12 in Appendix.



Thus, it appears that the unexpected negative effect of green party popularity seen initially, soon diminishes, and turns positive. This finding provides mixed support for H18 theorising a positive effect. As none of the theories assumes such a complex relationship between green party seats and compliance, they are unable to explain the change in direction. The results might support a theory of internalisation of norms, arguing that the longer one has been a member to the Protocol, the more the population legitimises it, and their increased environmental concern is materialised through support for Green parties, strengthening their position against the established parties (Bernhagen 2008). Yet, a look into how the variable varies with time reveals that Green seats numbers above 0 are evenly spread out throughout the research period, frustrating such an understanding. Even though the theories presented are unable to explain the observed connection, method-specific issues might be part of the reason. A descriptive analysis shows that only 218 country-years report green seats numbers above 0, suggesting that each unit's compliance record heavily influences the coefficient. Recall also that the operationalisation has been questioned, as it does not take into account the relative influence of the green parties expressed through the size of the legislative.

When included in the synthesis model, the Schoenfeld test no longer shows signs of non-proportionality for Green party popularity, thus leaving the original variable unchanged in the sample. Here, the predictor fails to reach significance, even though the hazard ratio reports the expected relationship. Thus, the covariate is not included in the final synthesis model. To sum up, the significant effect of one extra green seat in parliament is found to change from negative to positive in course of analysis time. The established theories can not explain this complex shape in hazard, and there is a chance that a skewed distribution may influence the hazard in a non-representative way. Thus, the results should be treated with care and further investigation should be conducted to determine the nature of this relationship.

*A left-libertarian executive* leads to a 20 percent point increase in the likelihood of complying with the Montreal Protocol, according to Model three. Even though the findings cohere with the theorised relationship, the variable is not significant, thus dismissing H19. Because of its relatively few observations, it is only included in one model. Furthermore, the descriptive statistics indicates that there is a clear bias towards developed countries. Looking closer at the data, it is made clear that the missing data are almost exclusively from Africa or the Middle East, regions where the party systems do not fit into Western party organisation and cleavage understandings. Thus, this measure would not have been interpreted with care even if it had come out as significant because of selection bias.

Lastly, *regional compliance* has throughout all models turned out to be significant at a .001-level, making it the best performer in this analysis. The logged variable tells us that a one percentage point positive change in the predictor leads to a 2,5 percentage point increase in the hazard, all other variables held constant and controlling for the Article 5-effect<sup>105</sup>. Confronting these numbers with empirical evidence, we expect Bolivia's hazard for experiencing compliance increases with 22,75 percent from 2005 to 2006 as its region moves from 0 to 9,1 percent compliance. The results confirm H20; the more compliant region, the more likely a member in this region is to move into compliance with the Montreal Protocol.

This positive effect is coherent with Simmons' (2000) findings in her study of commitment and compliance with the IMF. This finding is particularly intriguing as it has not, to my knowledge, been theorised in the literature on IEAs previously. This opens up great opportunities for focusing on compliance pull from smaller groups of countries through the sharing of experience within regions. Building on friendships and trust between neighbours, discussion and persuasion as well as technological advice from surrounding countries might be a good strategy for pushing member states towards compliance. Even though the framework of this thesis does not allow for generalisations, one is tempted to argue that such informal processes might be especially important within agreements lacking proper monitoring mechanisms. Moving one step down and focusing on compliance regionally, instead of centrally from the Ozone Secretariat, can possibly be an effective way of working towards maximising adherence to international environmental law.

## **6.6 Summary of the findings**

The synthesis models were constructed to provide a clear overview of which indicators that best explain the variation seen in compliance with the Montreal Protocol. Thus, only the six variables with reported significant effects in the same directions against several models were included in the synthesis models. The models showed us that aspects of all the IR theories, both the enforcement, managerial and legitimacy schools, are represented through the variables that are found to influence the likelihood of adhering to the set of obligations that the Montreal Protocol comprise. The final synthesis model ended up supporting four of the hypotheses set out in the theory chapter.

---

<sup>105</sup> Because  $2,452/100 = 0,0245$ , and  $\exp(0,0245) = 1,0248$ . The numbers are based on the final synthesis model.

**Table 9: The findings of the analysis**

	Variables	Expected effect	Actual effects				Implication of findings
			Model 1	Model 2	Model 3	Synthesis model(s)	
H1/H2	Country strength	Pos/Neg	<i>Negative</i>	<i>Negative</i>	<i>Negative</i>	<i>Negative</i>	Only H1 is supported
H3/H4	Economic integration	Pos/Neg	-	<i>Not sign</i>	<i>Not sign</i>	-	H3/H4 is rejected
H5	Government autonomy	Positive	-	<i>Not sign</i>	<i>Not sign</i>	-	H5 is rejected
H6	Government capacity	Positive	-	-	<i>Not sign</i>	-	H6 is rejected
H8	Democracy	Positive	-	<i>Not sign</i>	<i>Not sign</i>	-	H8 is rejected
H10	Economic development	Positive	-	<i>Positive</i>	-	<i>Positive</i>	H10 is supported
H11	French ccode/ Socialist law	Negative	<i>Not sign</i>	<i>Not sign</i>	<i>Not sign</i>	-	H11 is rejected
H13	MP-member length	Positive	<i>Positive</i>	<i>Positive</i>	<i>Positive</i>	<i>Positive</i>	H13 is supported
H15/H16	EU-member length	Pos/Neg	<i>Negative</i>	<i>Negative</i>	<i>Not sign</i>	<i>Not sign</i>	H15/H16 is rejected
H17	ENGO activity	Positive	-	<i>Not sign</i>	<i>Not sign</i>	-	H17 is rejected
H18	Green Party Popularity	Positive	<i>Neg/Pos</i>	<i>Neg/Pos</i>	<i>Neg/Pos</i>	<i>Not sign</i>	H18 is partly supported
H19	Left-libertarian executive	Positive	-	-	<i>Not sign</i>	-	H19 is rejected
H20	Regional compliance	Positive	<i>Positive</i>	<i>Positive</i>	<i>Positive</i>	<i>Positive</i>	H20 is supported

– indicates that the covariate was not included in the model. H7 Corruption, H9 Economic development squared, H12 Public opinion and H14 EU-membership were all excluded before the analysis was run.

The table shows that the hypotheses 1, 10, 13 and 20 are conclusively supported through this analysis. From this I conclude that a member state that has been a member to the Montreal Protocol for long, is surrounded by compliant neighbouring states and sees a high level of economic development has significantly better chances of experiencing compliance with the Montreal Protocol than a state not occupying these features. However, the state should not be too powerful (measured through country strength); this significantly decreases the chances of seeing compliance.

Turning to green party popularity and EU-membership, the analysis find mixed results. Green party popularity is found to be statistically significant in all three original models where it is included with an interaction term to even out its non-proportional effect on the hazard. When this interaction effect is excluded in the synthesis models, it no longer reaches significance. Still, there is a case to be made for the importance of green party popularity for compliant state behaviour. EU-membership, on the other hand, was rendered significant and

negative in Models 1 and 2 supporting H15 postulating that older EU-members have less chance of abiding to the emission standards. Yet it fails to reach a significant level in model 3 and the first synthesis model. Here as well, it can be that the effect is of importance, and that model specific issues have led to the reported non-significance.

Finally, there are seven variables and eleven hypotheses that are rendered no support in this analysis, and thus can be discarded with some confidence when considering the Montreal Protocol. Despite their insignificance, the direction of the relationship is inherently interesting and deserves a review. The variables *Left-libertarian executive* and *active ENGOs* are found to be positively related to the propensity of reaching compliance with the MP, corresponding with the theory. Only *Government capacity* is found to be negatively related to compliance, contradicting the postulations. The remaining variables *democracy*, *government autonomy* and *economic integration in world economy* change directions against different models. This may be a sign of badly fit data as their influence change depending on which covariates are included, or it may imply that the connection between these indicators and compliance is weak. Lastly, three of the four dummies checking for the effect of legal origin on compliance end up changing directions like the abovementioned variables. Only *French civil law* shows a positive effect on the compliance variable against all three models. Nonetheless, this finding contradicts the hypothesised relationship, and consequently is not rendered much trust.

## 7. Conclusion

The objective of this thesis has been to explore whether country characteristics can explain the observed variation in compliance with the Montreal Protocol. The motivation was offset by two features traced in the existing literature on compliance with environmental accords. First, the overwhelming explanatory power acclaimed to treaty-induced features without proper investigations of the influence of national mechanisms struck me as conspicuous. Secondly, the focus on overall compliance in the existing literature has led to a lack of understanding of which mechanisms facilitate compliance at the member-state level. By utilizing event history analysis I sought to figure out whether national structures and processes can reveal general patterns of influence on a member states' compliance record. This concluding chapter will first report the most important findings and use these to answer the research question put out for the thesis. Then, the theoretical implications of the findings will be discussed, before the thesis closes off with some suggestions for future research.

### 7.1 Explaining compliance with the Montreal Protocol

The findings from the analysis show that there is explanatory power to be attributed to factors external to the IEA framework. The analysis showed significant results across all models for four variables; *Country strength – population*, *Economic development*, *MP-membership – length* and *Regional compliance*. In addition, mixed support was found for two covariates; *EU-membership – length* and *Green party popularity*. In terms of overall model fit, the models in this analysis matter significantly, but their success in explaining the actual variation in compliance is moderate. The significant findings of length of treaty membership, regional compliance as well as for length of EU-membership indicate that there is a case to be made for the importance of learning effects. The longer a state is a member to the Protocol and to the EU, the more likely it is to legitimise the Montreal Protocol and internalise its norms. As contiguous countries achieve compliance, the non-compliant member is also 'caught in the undertow'. Also, the importance of power seems to hold as country strength clearly shows to reduce the likelihood of achieving compliance. Moreover, we see that external processes such as the member state's economy also play a significant role in explaining compliance. Lastly, the findings indicate that green party support may be a factor pushing a society towards compliance, but methodological issues as well as mixed results weakens this conclusion. Among the other covariates, it is surprising that the heavily theorised *ENGO activity* and *democracy* were not supported by the analysis. The compliance pull from international and

national environmental organisations thus can not explain the advance towards compliance seen in the Montreal Protocol context.

The above reported results allows me to answer the thesis' research question, concluding that *Country strength – population, Economic development, MP-membership – length and Regional compliance*, and to some extent, *EU-membership – length and Green party popularity*, are factors that explain member state compliance with the Montreal Protocol.

These findings add an important contribution to the established literature by pointing out that *one can not fully understand the difference seen in compliance with the Montreal Protocol without taking factors external to the environmental accords into concern*. The lack of systematic research on national factors has rightly been acknowledged by several IR-scholars, but despite this awareness few studies include country-specific variables in their research on the effect of IEAs on member states. An IR-scholar might contend against such an argument that national behaviour is understudied simply because it does not add anything to the study on overall compliance patterns. To that I would reply that one can not fully understand which mechanisms enable overall compliance without recognising the behaviour of the cornerstones in an international agreement, namely the signatory states. It is first when member state compliance is properly understood that effective and pertinent policy recommendations can be put out into life. Nonetheless, the findings of this thesis should not be considered a 'proof' against the literature focusing mainly of characteristics of the accords and the international environment when explaining compliance. The results add an important aspect to the existing literature focusing on institutional variables, but these two aspects are not mutually exclusive. To sum up, this thesis concludes that including state-specific processes next to more treaty-induced variables would provide the researcher with a more complete picture of the compliance process than could a study only including one group of variables.

## **7.2 Suggestions for further research**

When conducting research on a phenomenon, a goal is to be able to read policy recommendations from the inferences made. Yet, carving out new policy directions based on my study would at best be 'risky business,' as the scope of my study does not allow for generalisations. Nonetheless, further research on the mechanisms found to be significant here would be of great interest. If such studies should find similar patterns of influence, there might be a case to be made for an increased focus on social learning through more deliberate

channels, as well as a continued focus on economic as well as technological support.

Connected to this point is also the quest for more in-depth research on the mechanisms lying behind the variables that this study found to be of significant importance.

One of the great challenges with doing research on IEAs is their multifaceted nature. One cannot fully understand state behaviour in the international sphere without taking into account the history of the agreement, the protocol framework as well as each member state's qualifications to deal with the issue they have legally bound themselves to. The Montreal Protocol serves as a good example here. One important reason for its success was that the industry utilising ozone depleting substances had developed new technologies that facilitated a more smooth transition to a CFC and halon-free, but still competitively effective, industry. I was unable to include such a historical aspect in my statistical analysis even though I knew that it is important. This shows the need for continued qualitative research to get the whole picture. However, this does not depreciate the importance of statistical analysis as long as our units of research are states and our cases are environmental protocols.

When considering the complex nature of compliance processes, a mixed approach might be fruitful. Then, a natural next step could be to compare agreements with differing frameworks, to see how this affects compliance rates as well as the covariates under research. Are the same mechanisms important in these agreements? In that case, one might argue that there are some characteristics that seem to be more important than others for ensuring member state compliance with an agreement. If not, a discussion on whether one actually can compare country behaviour against different protocols might be necessary.

The research done on international environmental governance have been criticised for having great biases in their research. A main critique presented previously in this thesis is the problem of endogeneity; the fact that when considering regime outcomes, scholars tend only to analyse those who are already members to an agreement, rendering a positive bias in the findings when regime effects are considered. Thus, it would be interesting to test the effect of an agreement on a broader scale, for example through a panel analysis of reduction of pollution where also non-members are included. Interesting questions would be whether non-members also change their habits when an environmental accord is in place, and if so, how long it takes before the effects of an international framework trickles down to non-member states.

As already mentioned, the potential for generalisations about how states behave within the framework of international environmental agreements, is scarce in this analysis. The lack of quantifiable compliance information as well as the young age of most of today's existing

agreements makes quantitative research on a number of protocols hard. This should however become easier with time, as more information is gathered and statistical techniques are becoming more flexible. Thus, an important step towards a deeper understanding of state behaviour in the international realm is more quantitative research on the topic. The extensive qualitative research done has been important, but there is little doubt that supplies from a statistical approach will be substantial for developing a well-established paradigm on international environmental regimes. This should be a scholarly motivation, as understanding how members to an agreement act and what makes them comply is a key to securing our time's greatest challenge; saving the environment from man made destruction.



## Bibliography

- Adcock, Robert, and David Collier. 2001. "Measurement validity: A shared standard for qualitative and quantitative research." *American Political Science Review* 95 (3):529-46.
- Allison, Paul D. 1984. *Event history analysis: Regression for longitudinal event data*. 46 vols. California: SAGE Publications, Inc.
- Andresen, Steinar, Elin Lerum Boasson, and Geir Hønneland. 2008. *Internasjonal miljøpolitikk*. Bergen: Fagbokforlaget.
- Andresen, Steinar, and Lars H Gulbrandsen. 2003. "The role of green NGOs in promoting climate compliance."
- Axelrod, Regina S., David Leonard Downie, and Norman J. Vig. 2005. *The Global environment: institutions, law, and policy*. Washington: CQ Press.
- Barrett, Scott. 2003. *Environment and statecraft : the strategy of environmental treaty-making*. Oxford ; New York: Oxford University Press.
- Benedick, Richard E. 1991. *Ozone Diplomacy. New Directions in Safeguarding the Planet*. Cambridge, MA: Harvard University Press.
- Bernhagen, Patrick. 2008. "Business and international environmental agreements: domestic sources of participation and compliance by advanced industrialized democracies." *Global environmental politics* 8 (1):78-110.
- Binder, Seth, and Eric Neumayer. 2005. "Environmental pressure group strength and air pollution: An empirical analysis." *Ecological economics* 55 (4):527-38.
- Blossfeld, Hans-Peter, Alfred Hamerle, and Karl Ulrich Mayer. 1989. *Event history analysis : statistical theory and application in the social sciences*. Hillsdale, N.J.: L. Erlbaum Associates.
- Blossfeld, Hans-Peter, and Götz Rohwer. 1995. *Techniques of Event History Modeling: New Approaches to Causal Analysis*. Mahwah, New Jersey: Lawrence Erlbaum.
- Box-Steffensmeier, Janet M., and Bradford S. Jones. 1997. "Time is of the essence: Event history models in political science." *American Journal of Political Science* 41 (4):1414-61.
- . 2004. *Event history modeling : a guide for social scientists*. Cambridge ; New York: Cambridge University Press.
- Box-Steffensmeier, Janet M., Dan Reiter, and Christopher J. W. Zorn. 2003. "Nonproportional hazards and event history analysis in international relations." *The Journal of Conflict Resolution* 47 (1):33-53.
- Box-Steffensmeier, Janet M., and Christopher J. W. Zorn. 2001. "Duration models and proportional hazards in political science." *American Journal of Political Science* 45 (4):972-88.
- Breitmeier, Helmut, Oran R Young, and Michael Zürn. 2010. *The International Regimes Database*. Technische Hochschule Darmstadt 2009 [cited 30.03 2010]. Available from <http://www.fernuni-hagen.de/polis/lg2/projekte/InternationalRegimesDatabase.shtml>.
- Breitmeier, Helmut, Oran R. Young, and Michael Zürn. 2006. *Analyzing international environmental regimes : from case study to database*. Cambridge, Mass.: MIT Press.
- Brown Weiss, Edith, and Harold K. Jacobson. 1998. *Engaging countries : strengthening compliance with international accords*. Cambridge, Mass. ; London: MIT Press.
- Börzel, Tanja A, Meike Dudziak, Tobias Hofman, Diana Panke, and Carina Sprung. 2007. "Recalcitrance, inefficiency and support for European integration: Why member states do (not) comply with European law." In *Center for European studies Working Paper series*.

- Cameron, James, Jacob Werksman, and Peter Roderick. 1996. *Improving compliance with international environmental law*. London: Earthscan Pub.
- Chayes, Abram, and Antonia H. Chayes. 1993. "On compliance." *International organization* 47 (2):175-205.
- Chayes, Abram, and Antonia Handler Chayes. 1995. *The new sovereignty : compliance with international regulatory agreements*. Cambridge, Mass. ; London: Harvard University Press.
- Checkel, Jeffrey T. 2001. "Why comply? Social learning and European identity change." *International organization* 55 (3):553-88.
- CIA World Factbook. 2010. *Country comparison: population* 2010a [cited 18.03 2010]. Available from <https://www.cia.gov/library/publications/the-world-factbook/rankorder/2119rank.html?countryName=Antigua%20and%20Barbuda&countryCode=ac&regionCode=ca&rank=197#ac>.
- . 2010. *Field Listing: Legislative Branch*. Central Intelligence Agency 2010b [cited 17.09 2010]. Available from <https://www.cia.gov/library/publications/the-world-factbook/fields/2101.html?countryName=Bahrain&countryCode=ba&regionCode=me&#ba>.
- Cleves, Mario Alberto, Roberto Gutierrez, William Gould, and Yulia Marchenko. 2008. *An introduction to survival analysis using Stata*. 2nd ed. College Station, Texas: Stata Press.
- Congleton, Roger D. 1992. "Political institutions and pollution control." *The review of economics and statistics* 74 (3):412-21.
- Cox, David R. 1972. "Regression models and life-tables." *Journal of the Royal Statistical Society* 34 (2):187-220.
- Dahl, Robert. 1971. *Polyarchy*. New Haven: Yale University Press.
- Desai, Uday. 1998. "Environment, Economic Growth and Government." In *Ecological Policy and Politics in Developing Countries*, ed. U. Desai. Albany, NY: State University of New York Press.
- DeSombre, Elisabeth R. 2000. "Experience of the Montreal Protocol: Particularly Remarkable, and Remarkably Particular, The." *UCLA journal of Environmental Law and Policy* 19 (1):49-81.
- Doorenspleet, Renske. 2004. "The structural context of recent transitions to democracy." *European journal of political research* 43 (3):309-35.
- Downs, George W., David M. Roake, and Peter N. Barsoom. 1996. "Is the good news about compliance good news about cooperation?" *International organization* 50 (3):379-406.
- European Values Study Group and World Values Survey Association. 2006. "European and World Values Surveys four-wave integrated data file, 1981-2004."
- Geddes, Barbara. 1990. "How the cases you choose affect the answers you get: Selection bias in comparative politics." *Political Analysis* 2 (1):131-50.
- George, Alexander L., and Andrew Bennett. 2005. *Case studies and theory development in the social sciences*. Cambridge, Mass.: MIT Press.
- Gerring, John. 2004. "What is a Case Study and What is it Good for?" *The American Political Science Review* 98 (2):341-54.
- Giuliani, Marco. 2003. "Europeanization in Comparative Perspective: Institutional Fit and National Adaptation." In *The Politics of Europeanization*, ed. K. Featherstone and C. M. Radaelli. Oxford: Oxford University Press.
- Glaeser, Edward L., and Andrei Schleifer. 2002. "Legal Origins." *The Quarterly Journal of Economics* 117 (4):1193-229.
- Global Greens. 2010. *Green officeholders* 2010 [cited 11.05 2010]. Available from <http://www.globalgreens.org/officeholders>.

- Golub, Jonathan. 2008. "Survival analysis." In *The Oxford Handbook of Political Methodology*, ed. J. M. Box-Steffensmeier, H. E. Brady and D. Collier. Oxford: Oxford University Press.
- Grambsch, Patricia M., and Terry M. Therneau. 1994. "Proportional hazards tests and diagnostics based on weighted residuals." *Biometrika* 81 (3):515-26.
- Grant, Wyn, Duncan Matthews, and Newell. 2000. *The effectiveness of European Union environmental policy*. Basingstoke, UK: Macmillan.
- Grossman, Gene M., and Alan B. Krueger. 1995. "Economic Growth and the Environment." *The Quarterly Journal of Economics* 110 (2):353-77.
- Grønmo, Sigmund. 2004. *Samfunnsvitenskapelige metoder*. Bergen: Fagbokforlaget Vigmostad & Bjørke AS.
- Hadenius, Axel, and Jan Teorell. 2004. "Same, Same - but Different. Assessing alternative indices of democracy." In *2004 Annual Meeting of the American Political Science Association*.
- Hair, Joseph F., William C. Black, Barry J. Babin, Rolph E. Anderson, and Ronald L. Tatham. 2006. *Multivariate Data Analysis*. 6th ed. New Jersey: Pearson Prentice Hall.
- Hall, Peter A. 2003. "Aligning ontology and methodology in Comparative Research." In *Comparative historical analysis in the social sciences*, ed. J. Mahoney and D. Rueschemeyer. New York: Cambridge University Press.
- Harbaugh, William T., Arik Levinson, and David Molloy Wilson. 2002. "Reexamining the Empirical Evidence for an Environmental Kuznets Curve." *The review of economics and statistics* 84 (3):541-51.
- Haverland, Markus. 2000. "National adaptation to European integration: The importance of institutional veto points." *Journal of Public Policy* 20 (1):83.
- Honaker, James, and Gary King. 2010. "What To Do about Missing Values in Time-Series Cross-Section Data." *American Journal of Political Science* 54 (2):561-81.
- Hosmer, David W. 2002. "Review of An Introduction to Survival Analysis using Stata." *The Stata Journal* 2 (4):428-31.
- Hosmer, David W., and Stanley Lemeshow. 1999. *Applied Survival Analysis; regression modeling of time to event data*. New York: John Wiley & Sons, Inc.
- Hosmer, David W., Stanley Lemeshow, and Susanne May. 2008. *Applied survival analysis: regression modeling of time-to-event data*. Hoboken, New Jersey: Wiley-Interscience.
- Haas, Peter M. 1998. "Compliance with EU directives: insights from international relations and comparative politics." *Journal of European Public Policy* 5 (1):17-37.
- . 2008. *International environmental governance*. Aldershot, Hampshire, England ; Burlington, VT: Ashgate.
- Haas, Peter M., Robert O. Keohane, and Marc A. Levy. 1993. *Institutions for the earth: sources of effective international environmental protection*.
- Inglehart, Ronald. 1997. *Modernization and postmodernization: Cultural, economic, and political change in 43 societies*. Princeton: Princeton University Press.
- Inter-Parliamentary Union. 2010. *PARLINE database on national parliaments 2010* [cited 11.05 2010]. Available from <http://www.ipu.org/parline-e/parlinesearch.asp>.
- Jänicke, Martin. 1992. "Conditions for environmental policy success: an international comparison." *Environmentalist* 12 (1):47-58.
- Jänicke, Martin, and Helmut Weidner, eds. 1997. *National environmental policies. A comparative study of capacity-building*. Berlin: Springer.
- Kalbfleisch, John D., and Ross L. Prentice. 2002. *The Statistical Analysis of Failure Time Data*. Hoboken: John Wiley & sons, Inc.
- Kam, Cindy D., and Robert J. Franzese Jr. 2007. *Modeling and Interpreting Interactive Hypotheses in Regression Analysis*. Ann Arbor: University of Michigan Press.

- Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2009. "Governance Matters VIII. Aggregate and Individual Governance Indicators 1996-2008." The World Bank Development Research Group.
- Keech, William R. 1996. *Economic politics: The costs of democracy*. Cambridge: Cambridge University Press.
- Keefer, Phillip. 2009. "Database of Political Institutions: Changes and Variable Definition." World Bank Development Research Group.
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry : Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- Kuznets, Simon. 1955. "Economic Growth and Income Inequality." *The American Economic Review* 45 (1):1-28.
- La Porta, Rafael, Lopez-de-Silanes, Andrei Schleifer, and Rober Vishny. 1999. "The Quality of Government." *Journal of Law, Economics and Organization* 15 (1):222-79.
- Lafferty, W. M. William, and James Meadowcroft. 1996. *Democracy and the environment : problems and prospects / edited by William M. Lafferty and James Meadowcroft*. Cheltenham and Brookfield: Edward Elgar Publishing Limited.
- Lambsdorff, Johann Graf. 2006. "The Methodology of the TI Corruption Perceptions Index 2006." Transparency International and the University of Passau.
- Marshall, Monty G, and Keith Jagers. 2009a. "Polity IV Project, Dataset Version p4v2008.": Center for Global Policy.
- . 2009b. "Polity IV Project: Dataset Users' manual." Arlington: Center for Global Policy, George Mason University.
- McKeown, Timothy J. 1999. "Case studies and the statistical worldview: Review of King, Keohane, and Verba's *Designing social inquiry: Scientific inference in qualitative research*." *International organization* 53 (1):161-90.
- Meyer, John W., David J. Frank, Ann Hironaka, Evan Schofer, and Nancy B. Tuma. 1997. "The structuring of a world environmental regime, 1870–1990." *International organization* 51 (4):623-51.
- Midlarsky, Manus I. 1998. "Democracy and the Environment: An Empirical Assessment." *Journal of Peace Research* 35 (3):341-61.
- Midtbø, Tor. 2007. *Regresjonsanalyse for samfunnsvitere*. Oslo: Universitetsforlaget.
- Miles, Edward L., Arild Underdal, Steinar Andresen, Jørgen Wettestad, Jon Birger Skjærseth, and Elaine M. Carlin. 2002. *Environmental regime effectiveness : confronting theory with evidence*. Cambridge, Mass.: MIT Press.
- Mitchell, R. B. 2003. "International Environmental Agreements: A Survey of Their Features, Formation, and Effects." *Annual review of environment and resources* 28 (1):429-61.
- Mitchell, Ronald B. 1994. "Regime design matters: intentional oil pollution and treaty compliance." *International organization* 48 (3):425.
- . 1996. "Compliance theory: an overview." In *Improving Compliance with International Environmental Law*, ed. J. Cameron, J. Werksman and P. Roderick. London: Earthscan publishers.
- . 2004. "A Quantitative Approach to evaluating International Environmental Regimes." In *Regime Consequences. Methodological challenges and research strategies*, ed. A. Underdal and O. R. Young. Dordrecht: Kluwer Academic Publishers.
- Mooney, Christopher Z., and Mei-Hsien Lee. 2000. "The influence of values on consensus and contentious morality policy: US death penalty reform, 1956–82." *The Journal of politics* 62 (1):223-39.
- Moravcsik, Andrew. 1997. "Taking Preferences Seriously: A Liberal Theory of International Politics." *International organization* 51 (4):513-53.

- Munck, Gerardo L., and Jay Verkuilen. 2002. "Conceptualizing and measuring democracy: Evaluating alternative indices." *Comparative political studies* 35 (1):5-34.
- Neumayer, Eric. 2002. "Do democracies exhibit stronger international environmental commitment? A cross-country analysis." *Journal of Peace Research* 39 (2):139-64.
- . 2003. "Are left-wing party strength and corporatism good for the environment? Evidence from panel analysis of air pollution in OECD countries." *Ecological economics* 45 (2):203.
- Orr, John, Paul R. Sackett, and Cathy Z. Dubois. 1991. "Outlier detection and treatment in I/O psychology: a survey of researcher beliefs and an empirical illustration." *Personnel Psychology* 44 (3):473-86.
- Osborne, Jason. 2002. "Notes on the use of data transformations." *Practical assessment, Research and Evaluation* 8 (6).
- Panayotou, Theodore. 2003. "Economic growth and the environment." In *Economic survey of Europe: United Nations Economic Commission for Europe*.
- Péchoux, Isabelle, and Jérôme Pouyet. 2002. "Regulatory Competition, Transboundary Pollution and International Trade." In *Environmental Economics and the International Economy*, ed. L. Marsaliani. Dordrecht: Kluwer Academic Publishers.
- Pennings, Paul, Hans Keman, and Jan Kleinnijenhuis. 2006. *Doing Research in Political Science : An Introduction to Comparative Methods and Statistics*. 2nd ed. London: SAGE Publications.
- Perkins, Richard, and Eric Neumayer. 2007. "Implementing multilateral environmental agreements: an analysis of EU directives." *Global environmental politics* 7 (3):13-31.
- Political Risk Services Group, Inc. 2010. *International Country Risk Guide*. The PRS Group, Inc 2010 [cited 15.04 2010]. Available from <http://www.prsgroup.com/ICRG.aspx>.
- Putnam, Robert D. 1988. "Diplomacy and Domestic Politics: the Logic of Two-Level Games." *International organization* 42 (3):427-60.
- Ragin, Charles C. 2004. "Turning the tables: how case-oriented research challenges variable-oriented research " In *Rethinking social inquiry: diverse tools, shared standards*, ed. H. E. Brady and D. Collier. New York: Rowman and Littlefield.
- Ringquist, Evan J., and Tatiana Kostadinova. 2005. "Assessing the effectiveness of international environmental agreements: The case of the 1985 Helsinki Protocol." *American Journal of Political Science* 49 (1):86-102.
- Roberts, J. Timmons, Bradley C. Parks, and Alexis A. Vasquez. 2004. "Who ratifies environmental treaties and why? Institutionalism, structuralism and participation by 192 nations in 22 treaties." *Global environmental politics* 4 (3):22-64.
- Royston, Patrick, and D. G. Altman. 1994. "Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling (with discussion)." *Applied statistics* 43 (3):429-67.
- Royston, Patrick, and W. Sauerbrei. 2007. "Improving the robustness of fractional polynomial models by preliminary covariate transformation: A pragmatic approach." *Computational statistics & data analysis* 51 (9):4240-53.
- Runge, C. Ford. 1996. "Economic Trade and Environmental Protection." In *Environmental Policy with Political and Economic Integration*, ed. J. B. Braden, H. Folmer and T. S. Ulen. Cheltenham, UK: Edward Elgar.
- Sartori, Giovanni. 1970. "Concept misformation in Comparative Politics." *The American Political Science Review* 64 (4):1033-53.
- Sauerbrei, W., and P. Royston. 1999. "Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials." *Journal of the Royal Statistical Society. Series A. Statistics in society* 162 (1):71-94.



- Simmons, Beth A. 1998. "Compliance with international agreements." *Annual review of political science* 1 (1):75-93.
- . 2000. "International law and state behavior: commitment and compliance in international monetary affairs." *The American Political Science Review*:819-35.
- Skog, Ole-Jørgen. 2004. *Å forklare sosiale fenomener: en regresjonsbasert tilnærming*. 2nd ed. Oslo: Gyldendal akademisk.
- Sribney, William. 2010. *Likelihood-ratio test after survey/robust ML estimation*. STATA Data Analysis and Statistical Software 2005 [cited 30.09 2010]. Available from <http://www.stata.com/support/faqs/stat/lrtest.html>.
- Teorell, Jan, Nicholas Charron, Marcus Samanni, Sören Holmberg, and Bo Rothstein. 2009. "The Quality of Government Dataset." University of Gothenburg: The Quality of Government Institute.
- The Ozone Secretariat. 2010. *The 1987 Montreal Protocol on Substances that Deplete the Ozone Layer*. United Nations Environmental Programme 2004 [cited 17.03 2010]. Available from [http://ozone.unep.org/Ratification\\_status/montreal\\_protocol.shtml](http://ozone.unep.org/Ratification_status/montreal_protocol.shtml).
- The Ozone secretariat. 2010. *Status of ratification* 2009 [cited 18.03 2010]. Available from [http://ozone.unep.org/Ratification\\_status/](http://ozone.unep.org/Ratification_status/).
- . 2010. *Evolution of the Montreal Protocol*. United Nations Environmental Programme (UNEP), November 25th, 2009 2010 [cited 02.03 2010].
- The World Bank Group. 2010. *Governance Matters 2009. Worldwide Governance Indicators, 1996-2008*. The World Bank 2009a [cited 01.05 2010]. Available from <http://info.worldbank.org/governance/wgi/index.asp>.
- . 2010. *Quick query selected fro World Development Indicators* 2009b [cited 2010 24.03]. Available from <http://ddp-ext.worldbank.org/ext/DDPQQ/member.do?method=getMembers>.
- The World Databank. 2010. "World Development Indicators (WDI) and Global Development Finance (GDF)." The World Bank Group,.
- Treier, Shawn, and Simon Jackman. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* 52 (1):201-17.
- ul Haq, Mahbub. 2003. "The Human Development Paradigm." In *Readings in Human Development*, ed. F.-P. Sakiko, A. K. Shiva Kumar and A. Sen. New York: Oxford University Press.
- Underdal, Arild. 1998. "Explaining Compliance and Defection:: Three Models." *European Journal of International Relations* 4 (1):5-30.
- Underdal, Arild, and Oran R. Young. 2004. *Regime consequences : methodological challenges and research strategies*. Dordrecht: Kluwer Academic.
- United Nations. 2010. *World Macro Regions and Components*. United Nations 2000 [cited 04.05 2010]. Available from <http://www.un.org/depts/dhl/maplib/worldregions.htm>.
- United Nations Development Programme. 2002. "Human Development Report 2002: Deepening democracy in a fragmented world." United Nations.
- . 2010. *The Human Development Index (HDI)* 2009 [cited 24.03 2010]. Available from <http://hdr.undp.org/en/statistics/indices/hdi/>.
- United Nations Environmental Programme. 2010. *ODS Consumption in ODP Tonnes* 2010 [cited 17.03 2010]. Available from [http://ozone.unep.org/Data\\_Reporting/Data\\_Access/](http://ozone.unep.org/Data_Reporting/Data_Access/).
- United Nations Statistics Division. 2009. "National Accounts Main Aggregates Database." United Nations
- . 2010. *Millennium Development Goals Indicators* 2010 [cited 20.03 2010]. Available from <http://unstats.un.org/unsd/mdg/Metadata.aspx?IndicatorId=0&SeriesId=753>.

- Victor, David G., Kal Raustiala, and Eugene B. Skolnikoff. 1998. *The implementation and effectiveness of international environmental commitments : theory and practice*. Laxenburg: International Institute for Applied Systems Analysis; MIT Press.
- Vogel, David. 1997. *Trading up: Consumer and environmental regulation in a global economy*. Cambridge, MA: Harvard University Press.
- Yamaguchi, Kazuo. 1991. *Event history analysis*. Edited by B. Leonard and D. J. Rog. Vol. 28. Newbury Park: SAGE Publications.
- Young, Oran R. 1979. *Compliance and public authority: a theory with international applications*.
- . 1999. *The Effectiveness of International Environmental Regimes : Causal Connections and Behavioral Mechanisms* Cambridge: The MIT Press.

## APPENDICES

**Table 10: Countries included in the dataset**

Analyses of compliance with the Montreal Protocol	
Countries included in the analyses.	Afghanistan, <b>Albania</b> , Algeria, Angola, <b>Argentina</b> , Armenia, <b>Australia</b> , Austria, <b>Azerbaijan</b> , Bahrain, Bangladesh, <b>Belarus</b> , Benin, Bhutan, <b>Bolivia</b> , Bosnia and Herzegovina, <b>Botswana</b> , <b>Brazil</b> , <b>Bulgaria</b> , <b>Burkina Faso</b> , <b>Burundi</b> , <b>Cambodia</b> , Cameroon, <b>Canada</b> , <b>Central African Republic</b> , <b>Chad</b> , Chile, <b>China</b> , Colombia, <b>Comoros</b> , <b>Congo Brazzaville</b> , <b>Congo Kinshasa</b> , <b>Costa Rica</b> , <b>Croatia</b> , Cuba, <b>Cyprus</b> , <b>Czech Republic</b> , <b>Djibouti</b> , <b>Dominican Republic</b> , <b>Ecuador</b> , Egypt, <b>El Salvador</b> , <b>Equatorial Guinea</b> , <b>Eritrea</b> , <b>Estonia</b> , Ethiopia, <b>Fiji</b> , Finland, <b>France</b> , <b>Gabon</b> , <b>Gambia</b> , <b>Georgia</b> , <b>Germany</b> , <b>Ghana</b> , Greece, <b>Guatemala</b> , <b>Guinea-Bissau</b> , <b>Guinea</b> , <b>Guyana</b> , <b>Haiti</b> , Honduras, <b>Hungary</b> , <b>India</b> , <b>Indonesia</b> , <b>Iran</b> , <b>Israel</b> , <b>Italy</b> , <b>Ivory Coast</b> , <b>Jamaica</b> , <b>Japan</b> , <b>Jordan</b> , <b>Kazakhstan</b> , <b>Kenya</b> , Korea (north), Korea (south), Kuwait, Kyrgyzstan, <b>Laos</b> , <b>Latvia</b> , <b>Lebanon</b> , <b>Lesotho</b> , <b>Liberia</b> , <b>Libya</b> , <b>Lithuania</b> , <b>Macedonia</b> , <b>Madagascar</b> , <b>Malawi</b> , Malaysia, <b>Mali</b> , <b>Mauritania</b> , <b>Mauritius</b> , <b>Mexico</b> , <b>Moldova</b> , <b>Mongolia</b> , <b>Montenegro</b> , <b>Morocco</b> , <b>Mozambique</b> , <b>Myanmar (Burma)</b> , <b>Namibia</b> , <b>Nepal</b> , <b>The Netherlands</b> , <b>New Zealand</b> , <b>Nicaragua</b> , Niger, <b>Nigeria</b> , <b>Norway</b> , <b>Oman</b> , <b>Pakistan</b> , <b>Panama</b> , <b>Papua New Guinea</b> , Paraguay, <b>Peru</b> , <b>Philippines</b> , <b>Poland</b> , <b>Qatar</b> , <b>Romania</b> , <b>Russian Federation</b> , <b>Rwanda</b> , Saudi Arabia, Senegal, Serbia, <b>Sierra Leone</b> , <b>Singapore</b> , <b>Slovakia</b> , <b>Slovenia</b> , Somalia, <b>South Africa</b> , Spain, <b>Sri Lanka</b> , Sudan, <b>Swaziland</b> , <b>Sweden</b> , <b>Switzerland</b> , Syria, <b>Tajikistan</b> , Tanzania, <b>Thailand</b> , Togo, <b>Trinidad and Tobago</b> , <b>Tunisia</b> , <b>Turkey</b> , <b>Turkmenistan</b> , <b>Uganda</b> , <b>Ukraine</b> , United Arab Emirates, <b>United Kingdom</b> , <b>United States</b> , Uruguay, <b>Uzbekistan</b> , <b>Venezuela</b> , <b>Vietnam</b> , Yemen, <b>Zambia</b> , <b>Zimbabwe</b>
Font indicates compliance status as of end of 2008 for the Montreal Protocol:	
Regular:	
Non-compliant	
<b>Bold:</b>	
Compliant	
Countries excluded from the analysis because of less than 500 000 inhabitants	Antigua and Barbuda, Bahamas, Barbados, Belize, Brunei, Cape Verde, Cook Islands, Dominica, Grenada, Iceland, Kiribati, Liechtenstein, Luxembourg, Maldives, Malta, Marshall Islands, Federal states of Micronesia, Monaco, Nauru, Niue, Palau, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Samoa, Sao Tome and Principe, Seychelles, Suriname, Tonga, Tuvalu, Vanuatu, Vatican state
Left censored cases where compliance existed before ratification	Belgium, Denmark, Solomon Islands



**Table 11: Codebook**

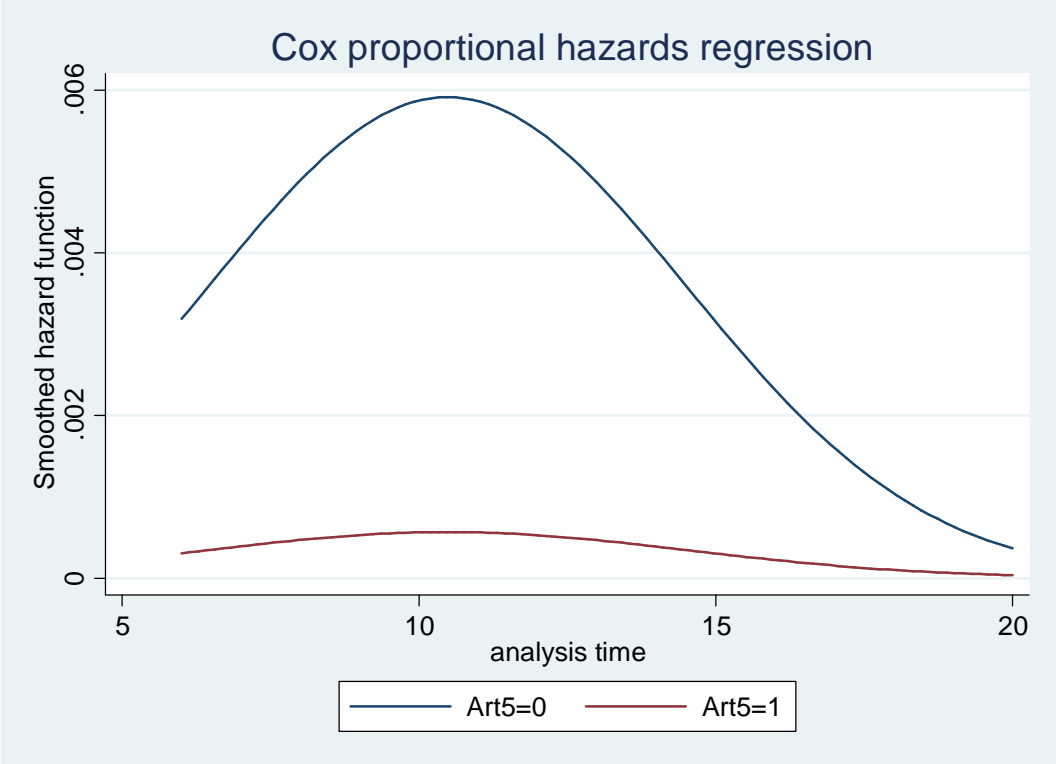
Variable	Source	Comments
Country strength	United Nations Statistics Division (2009), variable <i>population</i>	Reported into per million inhabitants to ease interpretation. Inversely squared.
Economic integration	The World Databank (2010), variable <i>merchandise trade as percentage of GDP</i>	Original measure divided by 10 to ease interpretation
Government autonomy	The World Bank Database of Political Institutions (Keefer 2009), variable <i>checks</i>	Based on an index of legislative and executive indices of electoral competitiveness
Government capacity	United Nations Development Programme (2009), <i>Human Development Index (HDI)</i>	Original measure is recoded to range from 0 to 100,
Corruption	The World Bank Worldwide Governance Indicators (WGI) (The World Bank Group 2009a), variable <i>control of corruption</i>	Only available 1996-2008
Democracy	Polity IV Project p4v2008 (Marshall and Jaggers 2009a), variable <i>polity2</i>	Range: -10 to +10
Economic development	The World Bank Group (2009b), <i>GDP per capita PPP, constant 2005 US dollars</i>	Logged. Reported per 1000 US dollars
Legal system	La Porta et al. (1999), gathered from (Teorell et al. 2009) variable <i>legal origin</i>	Four dummies (fifth as reference category) created based on the nominal variable <i>legal origin</i>
MP-membership	Based on information from the Ozone Secretariat (2009)	Coded by the author. Inversely squared.
EU-membership	Based on information from Perkins and Neumayer (2007)	
ENGO-activity	Yearbook of International Organization, reported by UNDP Human Development Report (2002), variable <i>Non-governmental organizations</i>	Constant variable, based on numbers from 2000. Reported per 100 NGO in a country.
Green party popularity	Based on Inter-Parliamentary Union's (2010) PARLINE database and Global Greens (2010)	Coded by the author
Left-libertarian executive	World Bank Database of Political Institutions (Keefer 2009), variable <i>execlrc</i>	Nominal variable <i>execlrc</i> recoded into dummy
Regional compliance	Based on information from the Ozone Secretariat (2009)	Coded by the author. Regions are defined using United Nations (2000). Logged.

**Table 12: Schoenfeld tests, run on all models**

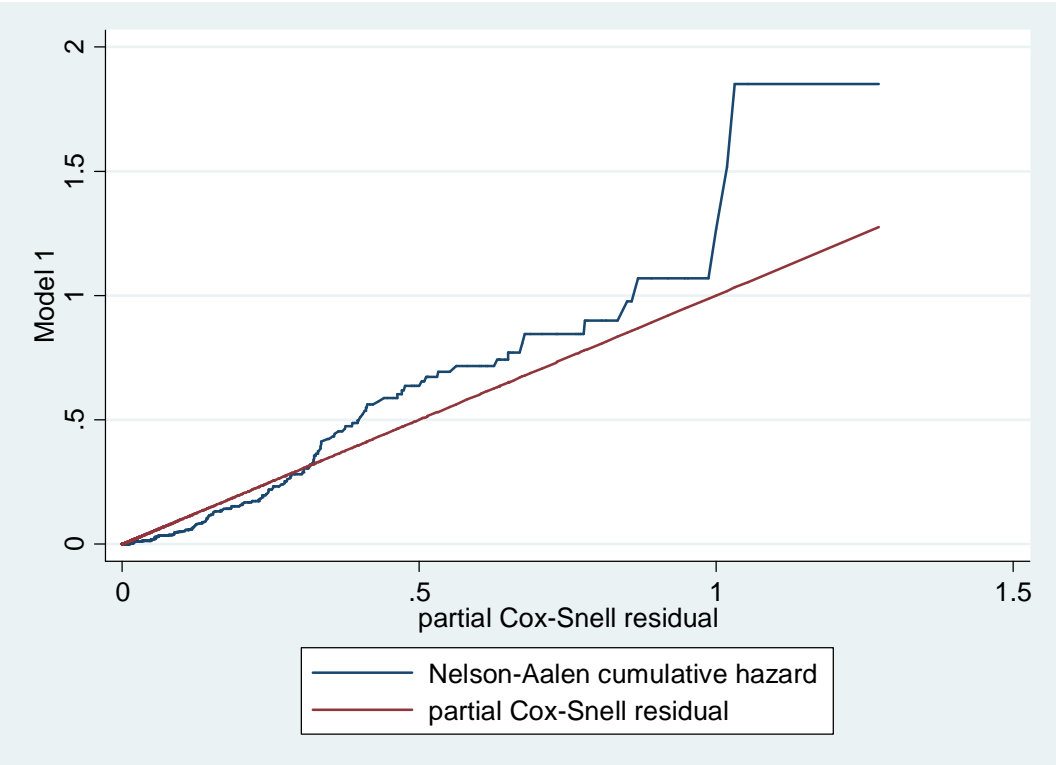
	Model 1	Model 2	Model 3	Synthesis M1	Synthesis M2
Country	0.076	0.050	0.052	0.053	0.044
strength	4.80*	0.90	0.56	2.54	1.67
C.strength inv. squared	0.056 0.26	0.047 0.26	0.079 0.67	0.072 0.47	0.043 0.17
Economic intergration		0.012 0.02	0.043 0.23		
Government autonomy		0.073 1.29	-0.013 0.03		
Government capacity			0.029 0.07		
Democracy		0.011 0.02	0.008 0.01		
Economic development		-0.011 0.02		-0.091 1.16	-0.061 0.48
Common law	0.032 0.17	0.025 0.10	0.046 0.21		
French civil law	0.060 0.59	0.053 0.44	0.047 0.20		
Socialist law	0.047 0.35	0.046 0.33	0.042 0.16		
German civil law	-0.083 1.55	-0.074 1.20	-0.137 1.69		
MP-memship length	0.066 0.44	0.060 0.51	0.095 1.15	0.075 0.64	0.070 0.56
MP-memship inv. squared	0.058 0.41	-0.015 0.03	0.047 0.16	0.102 1.29	0.099 1.15
EU-memship length	-0.098 0.86	-0.048 0.29	-0.081 0.55	-0.025 0.05	
NGO-activity		-0.016 0.04	0.013 0.02		
Green party popularity	0.167 3.50•	0.139 3.07•	0.221 3.86*	0.118 1.79	
Left-lib. executive			0.140 1.60		
Regional compliance	0.098 1.06	0.064 0.45	-0.026 0.05	0.060 0.36	0.066 0.42
Global test	10.97	11.62	14.22	3.59	1.11

The first row in each cell gives information about individual rho-values, chi squared values follow below.  
Significance: \*\*\* p<.001 \*\* p<.01 \* p<.05 • p<.10

**Figure 2: Baseline hazards for the two groups of Montreal Protocol members**



**Figure 3: Cox-Snell residuals, run on the biggest model (Model 1)**



Models the predictive power of the stratified Cox Model one. If the model fits the data well, the cumulative hazard of the Cox-Snell residuals should follow the 45 degree line (because then the residuals should have a standard exponential distribution with hazard functions that are equal to 1 for all  $t$  (Cleves et al. 2008:213-217). The overall fit is found to be good in this model. The variability in the right-hand tail is expected as prior failures and censoring leads to reduced effectiveness in the sample.

## Other tests run on the data

Residual based checks are run on the data to check for observations with a *disproportionate influence* on the estimated parameters. I use deviance residuals, meaning a rescaling of Martingale residuals to make them symmetric around zero, thus avoiding skewed data and hardened interpretation. It identified five countries with relatively large *dfbeta* values across several covariates. The question of how to deal with such outliers is heavily debated in the literature (Osborne 2002). Some argue that the only methodologically sound thing to do is exclude such observations if they turn out to be invalid, whereas others warn against such manipulation of the data (Orr et al. 1991). I run a sensitivity test, excluding the identified countries with influential observations from the sample and running regressions on Models 1 and 2. Even though most parameters looked the same, *green party popularity* is no longer jointly significant in Model 1 and 2. Additionally, *EU-membership – length* is no longer significant at the .10-level, and *regional compliance* reports a stronger effect on the hazard for experiencing compliance in Model 2. We see that there is an effect, albeit moderate, of excluding the outliers. Nonetheless, I argue for the inclusion of such influential observation in the dataset as there is no reason for assuming that these data are invalid.

A *sensitivity test* is run on the three original models, including left-censored countries that were excluded from the analysis. In this sample, the left-censored member states are coded as having complied with the Montreal Protocol in their first year of observation. The rerun of the analysis shows no real changes in Model 1. For Model 2, *EU-membership - length* just slips the .10-level, and thus is not longer considered significant, even though the effect was barely significant in the original model as well. Also, *German civil law* changes direction from positive in Model 2 to negative in the rerun model, but is still insignificant. As the specification of the legal origin dummies make them fragile for changes, this is not surprising. Model 3 still reports an opposite effect on *German civil law*, and *country strength* and *country strength inversely squared* just fail to reach the .10-level. All in all, these changes to the estimated predictors are considered to be small.

Another test named Harrell's C concordance statistics evaluates the *predictive power of the Cox model* on a member state's likelihood of reaching compliance. Unfortunately, I was not able to run this test in STATA because the test could not be used with late entry and time-varying data.

**Table 13: Pairwise correlations**

Sample: The Montreal Protocol	Dem .	Gov. aut.	Gov. cap.	Left-lib. exec	Green pop.	NGO pop.	EU-mem .	IEA-length	Reg. comp.	Ec. dev.	Ec. Int	Corr.	C. strength-gth	Com. law	Fr. ccode	Soc. law	Ger. ccode	Art.5
Democracy	1.00																	
Gov. aut.	0.50	1.00																
Gov. cap.	0.42	0.17	1.00															
Left-libexec	-0.27	-0.20	-0.12	1.00														
Green party	0.18	0.13	0.27	-0.01	1.00													
NGO	0.50	0.35	0.67	-0.10	0.48	1.00												
EU-length	0.22	0.14	0.36	-0.03	0.63	0.68	1.00											
IEA-length	0.22	0.10	0.30	-0.10	0.16	0.32	0.20	1.00										
Reg. comp.	0.29	0.18	0.44	0.05	0.22	0.47	0.26	0.42	1.00									
Ec. develop.	0.46	0.19	0.92	-0.14	0.29	0.73	0.39	0.32	0.45	1.00								
Ec. Int	-0.16	-0.17	-0.05	0.18	-0.01	-0.28	-0.02	0.05	0.18	-0.09	1.00							
Corruption	0.51	0.22	0.67	-0.15	0.33	0.76	0.43	0.27	0.50	0.78	-0.12	1.00						
C. strength	-0.05	0.23	0.03	0.05	0.18	0.46	0.21	0.19	0.04	0.01	-0.45	0.04	1.00					
Com.law	0.19	0.16	-0.13	-0.09	-0.13	-0.00	-0.12	0.15	0.12	-0.04	-0.08	0.19	-0.01	1.00				
Frenchlaw	-0.01	-0.50	-0.09	-0.09	-0.02	-0.04	0.15	0.03	-0.38	-0.06	-0.18	-0.17	0.03	-0.52	1.00			
Socialistlaw	-0.30	-0.16	0.07	0.22	-0.11	-0.19	-0.16	-0.27	0.19	-0.09	0.35	-0.24	-0.11	-0.35	-0.46	1.00		
Germanlaw	0.13	0.04	0.18	-0.44	0.45	0.26	0.23	0.08	0.06	0.24	-0.13	0.22	0.24	-0.13	-0.17	-0.12	1.00	
Art.5-mem	-0.30	-0.19	-0.63	0.04	-0.29	-0.66	-0.42	-0.13	-0.57	-0.62	-0.01	-0.63	-0.04	0.00	0.34	-0.24	-0.14	1.00

Correlation coefficients. Pairwise correlations are run on the largest sample from the multivariate analysis of the Montreal Protocol dataset.

**Table 14: Reported hazard ratios from Green party popularity at various points in time**

Analysis time	Model 1	Model 2	Model 3
$t = 1$	.933	.931	.933
$t = 2$	.946	.944	.949
$t = 5$	.987	.985	.999
$t = 6$	1.001	.999	1.017
$t = 7$	1.015	1.013	1.034
$t = 10$	1.059	1.057	1.089
$t = 15$	1.136	1.135	1.186
$t = 19$	1.201	1.201	1.270
(max value $t=20$ )			
Function changes			
direction at $t =$	6	6.1	5

Hazard rate for Greenseats at analysis time  $t = \exp[\beta_1 + \beta_2(t)]$ , where  $\beta_1$  is the coefficient for *Greenseats* and  $\beta_2$  is the coefficient for *Greenseats\*analysis time* (Hosmer et al. 2008:116-120). Hazard ratios is interpreted relative to one, lower values indicating a negative relationship and values higher than 1 implying positive connection.

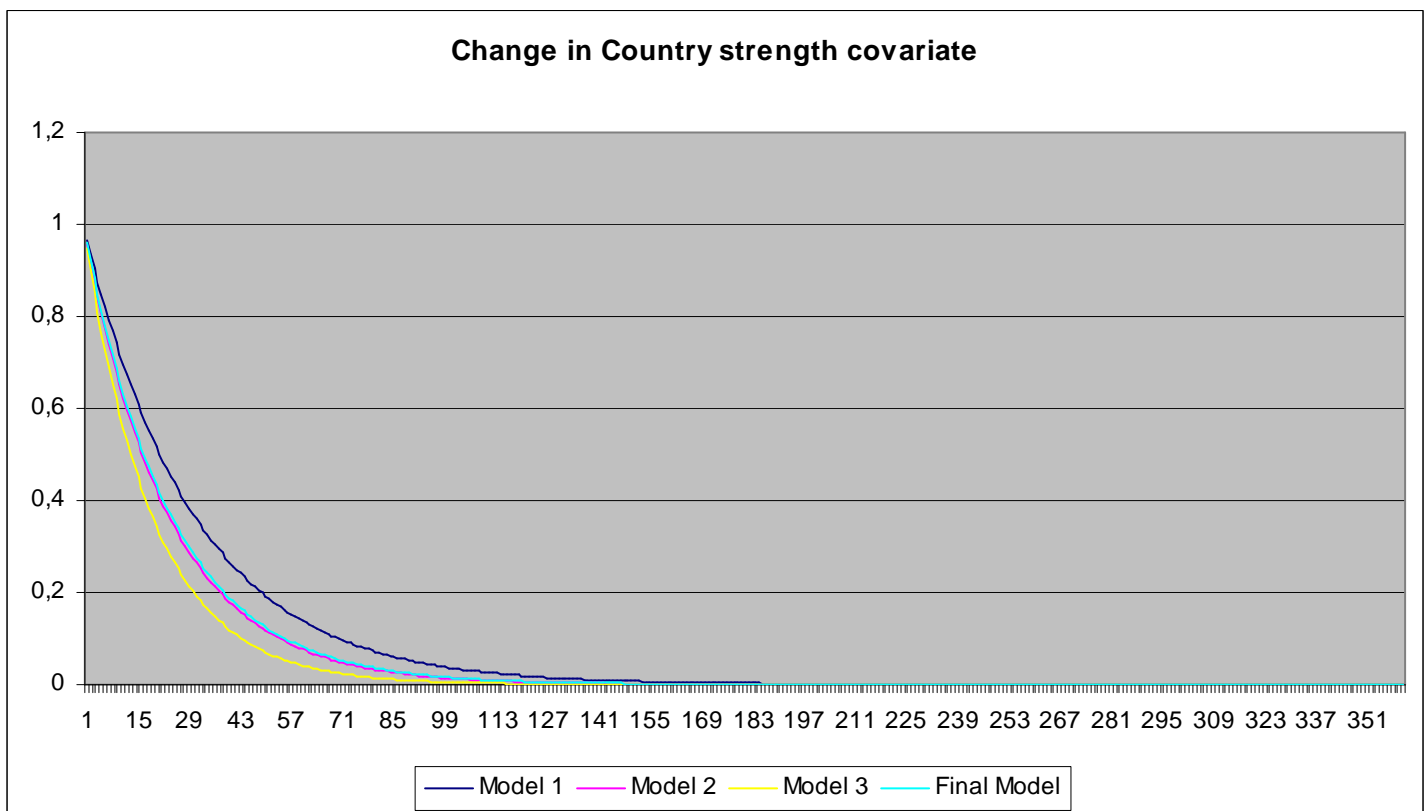
**Table 15: Change in hazard ratio from a one million increase in population**

1 million added at:	Model 1	Model 2	Model 3	Final Model
1 million	.905	.879	.853	.882
10 million	.500	.406	.328	.417
30 million	.133	.073	.039	.079
50 million	.040	.013	.005	.015
70 million	.010	.002	.001	.003
90 million	.003	.000	.000	.001
150 million	.000	.000	.000	.000

(max value 1314 mill.)

HR =  $\exp[\beta_1 c + \beta_2 (2 * Cstrength * c + c^2)]$  where  $\beta_1$  is the coefficient for *Country strength*,  $\beta_2$  is the coefficient for *Country strength inversely squared* and c is the change in Country strength at certain values of the variable, in this case one year (Hosmer and Lemeshow 1999:136). After having found the hazard rate for a squared function, the results are inversed  $\left(\frac{1}{HR}\right)$  to reflect the interaction.

**Figure 4: The shape of the hazard for the different models, based on Table 15**



**Table 16: Change in hazard ratio from one additional year as member of the Montreal Protocol**

1 year added at:	Model 1	Model 2	Model 3	Final Model
0 member years	1.090	1.103	1.111	1.057
1 member years	1.167	1.186	1.228	1.127
5 member years	1.431	1.423	1.658	1.367
10 member years	2.012	1.747	2.736	1.886
15 member years	2.829	1.957	4.516	2.602
20 member years (max observed)	3.978	2.056	7.452	3.589

$HR = \exp[\beta_1 c + \beta_2 (2 * MP - length * c + c^2)]$  where  $\beta_1$  is the coefficient for *MP-length*,  $\beta_2$  is the coefficient for *MP-length inversely squared* and  $c$  is the change in *MP-length* at certain values of the variable, in this case one year (Hosmer and Lemeshow 1999:136). After having found the hazard rate for a squared function, the results are inversed  $\left(\frac{1}{HR}\right)$  to reflect the interaction.

**Figure 5: The shape of the hazard for the different models, based on Table 16**

